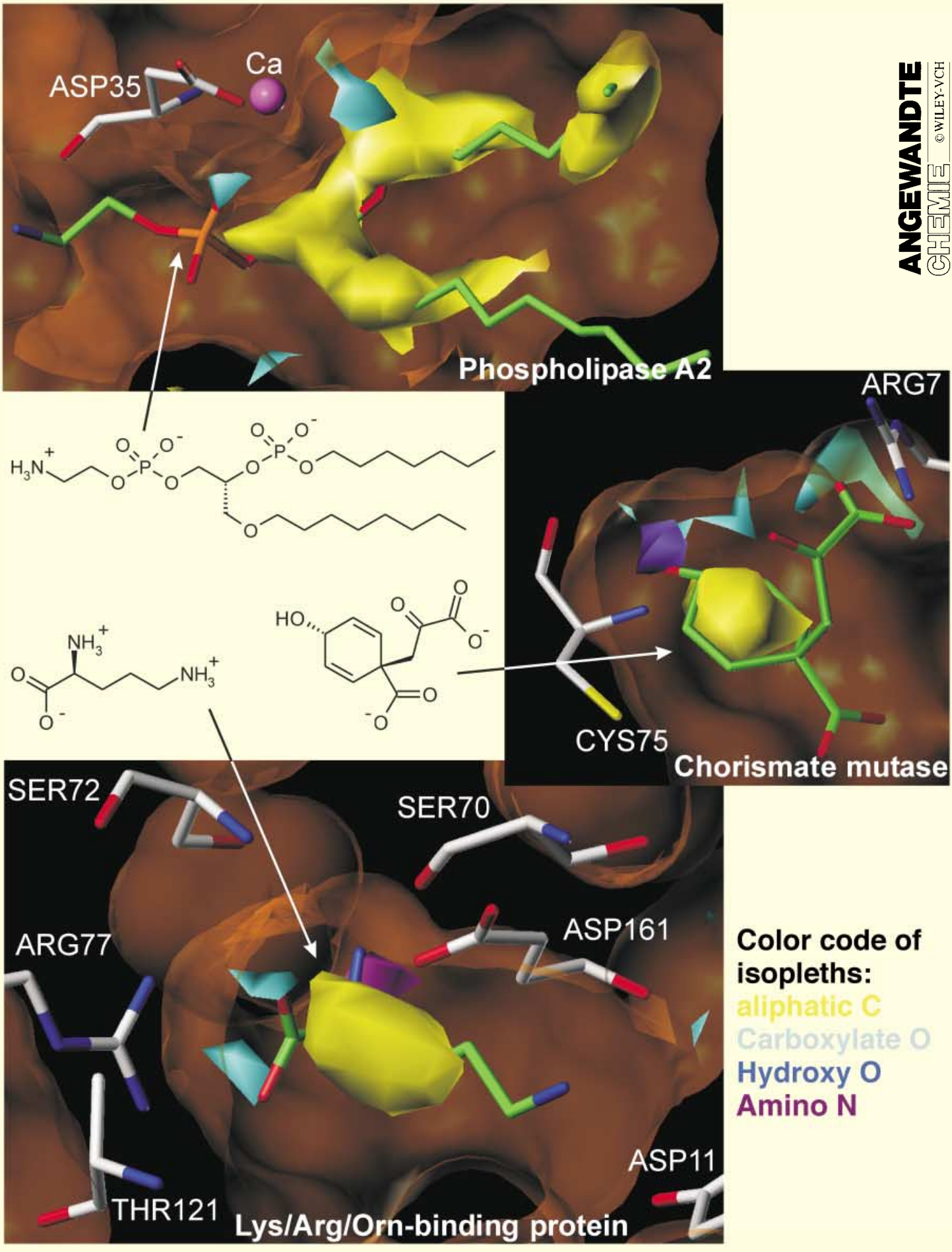
Spatially resolved identification of favorable protein–ligand interactions in binding pockets



ANGEWANDTE CHEMIE

Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors

Holger Gohlke and Gerhard Klebe*

The influence of a xenobiotic compound on an organism is usually summarized by the expression *biological activity*. If a controlled, therapeutically relevant, and regulatory action is observed the compound has potential as a drug, otherwise its toxicity on the biological system is of interest. However, what do we understand by the biological activity? In principle, the overall effect on an organism has to be considered. However, because of the complexity of the interrelated processes involved, as a simplification primarily the "main action" on the organism is taken into consideration. On the molecular level, biological activity corresponds to the binding of a (low-molecular weight) compound to a macromolecular receptor, usually a protein. Enzymatic reactions or signal-transduction cascades are thereby influenced with respect to their function for the organism. We regard this binding as a process under equilibrium conditions; thus, binding can be described as an association or dissociation process. Accordingly, *biological activity* is expressed as the *affinity* of both partners for each other, as a thermodynamic equilibrium quantity. How well do we understand these terms and how well are they theoretically predictable today? The holy grail of rational drug design is the prediction of the biological activity of a compound. The processes involving ligand binding are extremely complicated, both ligand and protein are flexible molecules, and the energy inventory between the bound and unbound states must be considered in aqueous solution. How sophisticated and reliable are our experimental approaches to obtaining the necessary insight? The present review summarizes our current understanding of the binding affinity of a small-molecule ligand to a protein. Both theoretical and empirical approaches for predicting binding affinity, starting from the three-dimensional structure of a protein–ligand complex, will be described and compared. Experimental methods, primarily microcalorimetry, will be discussed. As a perspective, our own knowledge-based approach towards affinity prediction and experimental data on factorizing binding contributions to protein–ligand binding will be presented.

Keywords: binding affinity · calorimetry · drug research · protein–ligand interactions · scoring function

1. Introduction

Mutual molecular recognition is the starting point for almost all processes in biological systems. More than 100 years ago, this fact was first recognized by Emil Fischer, who wrote "*that enzyme and glycoside must fit together like a key and a lock in order to initiate a chemical action upon each other*".[3] Also Paul Ehrlich's statement *"Corpora non agunt nisi fixata"*[**][4] is—in a somewhat extended form[5, 6]—the basis for the scientific explanation of drug action. As a consequence, the geometrical and chemical complementarity of small molecules (termed *ligands* in the following) and their macromolecular biological target structures (mostly proteins, termed *receptors* in the following) influences metabolic or signal-transduction pathways, and thus initiates a physiological effect.

In recent years, knowledge of the relationship between molecular structure and biological effects has prompted a fundamental change of the methods used in modern drug research. Molecular biological techniques identify receptor dysfunction or failures in regulation as possible causes of a disease. Furthermore, they help to isolate proteins in a purified form, the three-dimensional structure of which is subsequently determined using X-ray structural analysis,[7, 8] NMR spectroscopy,[9–11] or cryoelectron microscopy.[12–15] In addition, the number of characterized protein sequences is currently growing at a dramatic rate as a result of several genome-sequencing projects.[16–19] This provides a platform for

[*] Prof. Dr. G. Klebe, Dr. H. Gohlke
Institut für Pharmazeutische Chemie
Philipps-Universität Marburg
Marbacher Weg 6, 35032 Marburg (Germany)
Fax: (+49) 6421-282-8994
E-mail: klebe@mailer.uni-marburg.de

[**] "The bodies do not act if they are not bound."

  1433-7851/02/4115-2645 $ 20.00+.50/0

the development of improved techniques to predict protein function[20–24] and their three-dimensional structure.[25, 26] Furthermore, enhanced methods for crystal-structure analysis[27–30] are currently being developed and computational procedures are being applied to identify and select new biological targets.[31–33] As a result of these efforts, in the near future an enormously expanded range of structurally and functionally characterized molecular targets will become available for drug therapy.[34]

The development of a new drug, which can take up to 15 years and consume about half a billion dollars,[35–37] can be divided into several phases.[38, 39] The process starts with an initial search for a lead structure, that is, a ligand with a detectable affinity for a given receptor. This step is followed by several stages of optimization. The increase in affinity and selectivity of a ligand towards its biological target must be accompanied by an optimization of beneficial pharmacokinetic properties. This involves the absorption of the drug, its distribution and metabolism in the body, along with its excretion and toxicity.[40–42] The actual validation of a drug is subsequently performed in several phases of clinical trials.

The rapid and reliable identification of potent, high-affinity ligands is of utmost importance in view of the overwhelming number of characterized receptors expected from the genome projects.[43, 44] Given the limited resources available, the proof of concept for the relevance of a particular therapeutic target has to be assessed in the early phase of drug development. Presently, there are two complementary approaches in the search for new lead structures: *experimental* (high-throughput) *screening*,[45, 46] involving the in vitro testing of large compound libraries, and *virtual screening*[47, 48] or *rational design*,[49–51] which is based on available information about the structure of the biological target and/or already characterized ligands.

Experimental random screening originated from methods of traditional drug research. Large compound libraries of synthetic and natural compounds are tested for possible activity in a bioassay, independent of their actual chemical structure.[52, 53] In recent years, this method has been promoted extensively by the use of robotic systems to achieve high-throughput testing[54] along with methodological developments towards combinatorial chemistry,[55–59] and automated parallel synthesis. Using these techniques, libraries with several tens of thousands of compounds can be easily synthesized in a short time starting from a few reagents. However, the hit rates obtained by these time- and cost-intensive methods, also termed "irrational" because of their untargeted character ("as many as possible and as rapidly as possible"), are frequently less than one tenth of a percent of the number of compounds tested.[60, 61] Moreover, this method usually ignores knowledge of the features of the biological target. As a consequence, nonrandom or targeted approaches to screening have been developed. Here, the test compounds are preselected by computer methods to maximize their pairwise diversity, for example with respect to chemical properties,[62] the expected favorable pharmacokinetic behavior,[63, 64] or the biological target molecule.[65–69]

Rational drug design follows a different approach. Starting from a known or hypothetical mode of action or binding mechanism, a lead structure is rationally designed and subsequently tested experimentally. The obtained results are fed back into a design cycle as new information (Figure 1).[51, 70] Impressive results have been obtained with this strategy, as presented, for example, in the recently published studies on the discovery of inhibitors of DNA gyrase[71] or matrix metalloproteinase 13.[72] Although this approach is still in its infancy and has only recently profited from advances in computer technology and methodology,[73, 74] there are already

Holger Gohlke, born 1972, studied chemistry in Darmstadt and completed his Diplomarbeit in 1997 under the supervision of Prof. F. W. Lichtenthaler. He obtained his doctorate in 2000 with Prof. G. Klebe, Philipps-Universität Marburg, where he developed methods for the prediction of protein–ligand interactions. Currently, he is working as a Feodor Lynen Research Fellow with Prof. D. A. Case at the Scripps Research Institute, La Jolla, USA. His research interests are in the area of computational chemistry and computer-aided drug design.


H. Gohlke


G. Klebe

Gerhard Klebe, born 1954, studied chemistry at the University of Frankfurt and obtained his doctorate in physical chemistry. As a Studienstiftung scholar he spent a year in Grenoble, France, at the CNRS and ILL working on experimental electron-density determinations. After postdoctoral positions in crystallography (H. Fuess and H. B. Bürgi), he joined BASF-AG in Ludwigshafen, where he was responsible for molecular modeling and crystallography in drug research. He obtained his Habilitation in pharmaceutical and structural chemistry at the University of Heidelberg in 1992. In the summer of 1996, he was appointed full Professor for Pharmaceutical Chemistry at the University of Marburg. The focus of his research is directed towards various aspects of drug design, in particular the interaction of small-molecule ligands with proteins. The presently applied methods involve molecular biology and protein chemistry, crystallography, bioinformatics and software development, application of drug-design methods, microcalorimetry, and the synthesis of initial lead structures.

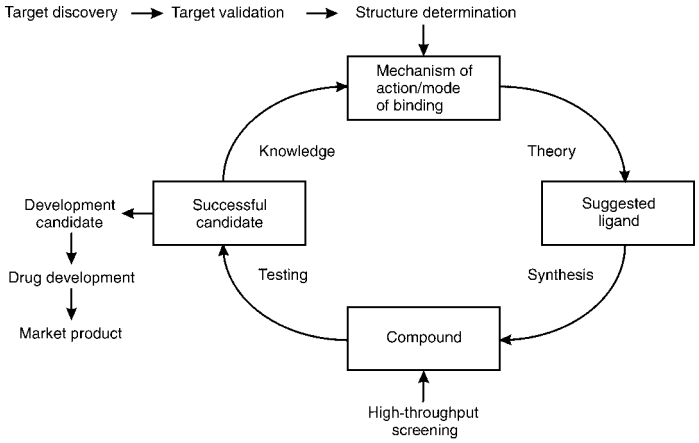
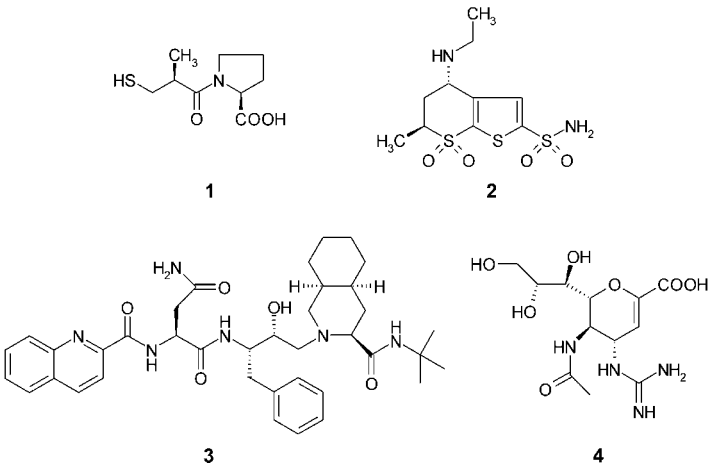


Figure 1. General approach for the rational design of inhibitors. Starting with discovered or previously synthesized compounds and biological testing, information about the mechanism of action or binding mode is used as a starting point for a subsequent design cycle assisted by computational methods.



Scheme 1. Drugs used in therapy, whose development was significantly supported by rational design.

a fair number of examples in which the development and optimization of drug candidates have strongly profited from this approach[75–79] (Table 1).

Table 1. Proteins, for which inhibitors were discovered or optimized by rational drug design.

Protein	References
aldose reductase	[80, 81]
calmodulin	[82]
carboanhydrase	[83–86]
cyclooxygenase-2	[87, 88]
elastase	[89–91]
FKBP12	[92, 93]
gyrase	[94]
HIV protease	[95–103]
papain	[104]
purine nucleoside phosphatase	[105]
renin	[106–109]
reverse transcriptase	[110–117]
selectin	[118–121]
sialidase (neuraminidase)	[122–132]
streptavidin	[133–136]
thermolysin	[137]
thrombin	[138–145]
thymidylate synthase	[146]

Furthermore, a number of drugs have already been introduced into therapy, which were discovered by this strategy, or where rational design has played a key role in the discovery process, for example, the angiotensin-converting-enzyme (ACE) inhibitor captopril (**1**),[147] the carboanhydrase inhibitor dorzolamide (**2**),[77] the HIV-protease inhibitors saquinavir (**3**), indinavir, ritonavir, and nelfinavir,[148] and the sialidase inhibitors zanamivir (**4**) and oseltamivir (Scheme 1).[149]

The strategy to be followed in rational design depends on whether the three-dimensional structure of the biological target molecule is known or not. In the latter case, "quantitative structure–activity relationships" (QSAR methods)[150–154] can be used to establish a relationship between molecular structure and biological activity for a series of active compounds. These models do not only explain the relative differences among the observed affinities, but also allow an affinity prediction for unknown compounds. An alternative procedure is the generation of a pharmacophore model from a series of active compounds.[155] Here, the molecular properties of the active compounds are represented in geometric terms, which are a prerequisite for biological activity. In a subsequent step new, potentially active, candidate molecules are retrieved from a compound library that obey the pharmacophore hypothesis.[156] In addition, the results from QSAR methods provide some insight into the structural requirements of the receptor responsible for the derived structure–activity relationship. This information can also be used to generate mini- or pseudoreceptor models.[157–160]

The gradually increasing number of structurally characterized macromolecular receptors[2] (Figure 2) provides the basis for any *structure-based design* of active compounds. Receptor geometries are predominantly determined by crystal-structure analysis. The obtained geometries are assumed to be relevant also for the conditions in solution.[161–166] Using information about the properties of the ligand-binding site along with the assumption, based on the lock-and-key

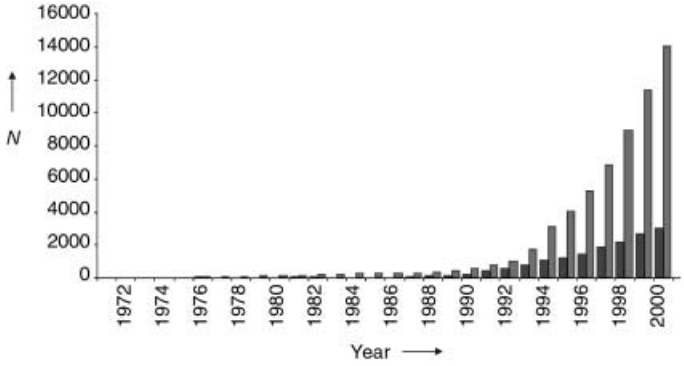


Figure 2. Total number of entries *N* stored in the protein databank PDB[2] (light gray bars) and entries newly deposited every year (dark gray bars; status January, 2001).

principle, that a potent ligand must exhibit significant structural and chemical complementarity with the binding pocket, two strategies for computer-aided drug design are possible: in de novo design, novel leads are generated in the binding pocket starting from prepositioned seed atoms or fragments that are subsequently grown into entire molecules.[167, 168] Alternatively, a compound library can be screened for ligands in agreement with the binding-site requirements. The individual molecules are flexibly docked into the binding site.[169] In both cases, a fast *prediction of ligand affinity* towards the receptor is the most crucial step: only if this assessment is performed with acceptable accuracy and reliability can new leads be discovered by computational methods. This latter strategy is known as virtual screening.[47, 48, 170] Besides considerably reducing the time and costs of the development, more importantly, the structural insights and affinity information learned can be used subsequently for lead *optimization*.

Two aspects determine the success of computer-aided structure-based ligand design: the generation of reasonable ligand-binding modes (*configuration-generation problem*) and the recognition of those binding modes that correspond best to the experimentally given situation based on a reasonable estimate of the expected binding affinity (*structure and affinity prediction problem*).[171] An objective test of docking methods performed in 1997[172] confirmed the assumption that recognizing near-native geometries and predicting their affinities could be achieved only with limited success,[173–175] whereas the problem of generating reasonable ligand orientations is considered to be virtually resolved,[50, 176] at least for proteins with rather rigid binding pockets, not involving any water molecules in binding and without any change in protonation state of either ligand or protein upon binding.[177]

In this review, our current knowledge of the description and prediction of binding affinities of small-molecule ligands to proteins will be summarized. In addition, we will present some of our own recent results on this topic which, on the one hand, give a better understanding of the thermodynamic aspects of ligand binding and, on the other hand, provide a clear improvement in computer docking and virtual screening.

2. Binding Affinity as a Result of Inter- and Intramolecular Contributions—Macroscopic Effects Resulting from Microscopic Events

2.1. 3D Receptor–Ligand Structures—Windows to the World of Interactions

3D structures of receptors (and ligands) do not only form the basis for structure-based drug design. Many of the aspects of ligand-binding described in the following would still be only superficially understood if we did not have the facilities to study them on a molecular level. Protein crystallography,[7, 8, 178] complemented by high-resolution NMR spectroscopy,[9–11] provides us with the required information about the arrangement of the atoms in space.

With respect to the study of protein–ligand complexes, it must be kept in mind that X-ray diffraction hardly differentiates between isotopes and elements of similar atomic number because of their comparable diffraction power. Except for protein structures at very high resolution, the positions of terminal N and O atoms, for example, in asparagine and glutamine, can only be assigned on the basis of a self-consistent hydrogen-bond network. Similar problems occur for the imidazole ring of histidine which can adopt two virtually indistinguishable orientations. Moreover, the positions of the poorly diffracting hydrogen atoms remain undetermined. This is of particular importance in the case of H atoms of conformationally flexible groups (for example, hydroxy or amino groups), as well as for groups that can be (de)protonated. However, a consistent picture of the localization of (polar) H atoms and the adopted protonation state can usually be assigned by analyzing the atoms in the neighborhood.[179, 180]

The relevance of the protein geometry obtained by crystal-structure analysis is mainly determined by the quality of the studied crystals. Resolutions below 1.5 Å (coinciding with the mean length of a covalent bond) are rarely obtained for protein crystals; values between 2 and 3 Å are more usual. Atomic resolution is achieved below 1.2 Å. Along with techniques for experimental phase determination, insights into the electronic structure of molecules[181,182] and the localization of polar H atoms can be achieved. The estimated standard-deviation value in atomic coordinates is inherently related to resolution, for example, for a resolution of 2.5 Å this value is about 0.4 Å.[183–185] This has to be considered whenever intermolecular interactions are discussed on the basis of protein crystal structures.

Furthermore, crystallography performs an averaging in time and space of the individual molecules forming the crystals.[163] Crystal contacts between neighboring molecules can result in intermolecular interactions, which may affect parts of a structure. Besides positional disorder, which results in distinct occupancies of alternative atomic positions, dynamic disorder also results from thermal motion of the atoms about their equilibrium positions. Even for atoms with an average Debye–Waller factor (B factor) of 20 $Å^2$, the mean atomic displacement from equilibrium amounts to 0.5 Å. Moreover, because of time and space averaging during data collection and the requirement for an unperturbed periodic arrangement, only spatially restricted atoms contribute constructively to diffraction and thus only their positions can subsequently be located.

This applies especially to water molecules which can make up to 70% of the number of atoms in a protein crystal.[186, 187] While the water molecules of the first hydration shell surrounding protein or ligand are generally well-ordered, their mobility increases with their distance from the molecular surface.[188]

Moreover, multiple binding modes of ligands can occur as a result of spatial averaging. In such situations, the same ligand can occupy several energetically equivalent orientations in the binding pocket. This case is especially true for weakly binding ligands. In addition, deviating binding modes can occur in different polymorphic forms of the crystalline state. Under

kinetically controlled conditions[190–192] deviating packing arrangements can be formed, which result in different crystal structures of distinct physicochemical properties.[189] In the case of crystals of proteins or protein–ligand complexes, these "polymorphs" are usually referred to as "different crystal forms". Nevertheless, often they exhibit deviating (enzymatic) properties; for example, depending on the crystallization conditions, lipase crystals can be obtained in an "open" or "closed" form.[166]

In summary, any referral to "the" crystal structure of a compound or protein–ligand complex has to be regarded with some care in the light of these effects that could lead to multiple crystal forms.

2.2. Factors Determining Ligand–Receptor Binding Affinity

The selective binding of a small-molecule ligand to a particular protein is determined by a mutual structural and energetic recognition.[193–195] Ligands can interact either covalently or noncovalently with their biological target.

The noncovalent, reversible association of receptor (R) and ligand (L) to form a receptor–ligand complex (R′L′) generally occurs in an aqueous, electrolyte-containing solution [Eq. (1)].

$$R_{aq.} + L_{aq.} \rightleftarrows R'L'_{aq.'} \quad (1)$$

Under thermodynamic equilibrium conditions, this reaction is determined by the standard Gibb's free energy (or free enthalpy, used in the following) of binding ΔG°. This quantity is related to the experimentally determined association constant K_A (or its reciprocal dissociation or inhibition constants, K_D or K_i, respectively) [Eq. (2)]. ΔG° is composed of an enthalpic (ΔH°) and an entropic ($T\Delta S^\circ$) portion. T refers to the absolute temperature.[196, 197] In place of ΔG°, the term (binding) affinity is used to describe the tendency of a molecule to form a complex with another one.

$$\begin{aligned} K_A &= K_D^{-1} = K_i^{-1} = \frac{[R'L']}{[R]\,[L]} \\ \Delta G^\circ &= -RT\ln K_A = \Delta H^\circ - T\Delta S^\circ \end{aligned} \quad (2)$$

According to Equation (3), with μ_i^0 as the chemical standard potential of the species i, ΔG° can also be understood as a function to describe the stability of the complex with respect to free ligand and uncomplexed receptor.[198, 199]

$$\Delta G^\circ = \mu^\circ_{R'L'aq.'} - (\mu^\circ_{Raq.} + \mu^\circ_{Laq.}) \quad (3)$$

Experimentally determined inhibition constants fall into a range between 10^{-2} and 10^{-12} M, which corresponds to a Gibbs free standard enthalpy of binding of -10 to -70 kJ mol^{-1} at $T = 298$ K.[194] A change in free enthalpy of 5.7 kJ mol^{-1} at this temperature alters the inhibition constant by one order of magnitude. A comparison of affinities of reversibly binding ligands shows that the increase in binding affinity is about 6.3 kJ mol^{-1} per atom for molecules with up to 15 non-hydrogen atoms. Moreover, for larger ligands the affinity is only slightly dependent on the molecular weight.[200]

It is generally accepted that electrostatic interactions determine noncovalent ligand–receptor binding. They comprise salt bridges, hydrogen bonds, dipole–dipole interactions, and interactions with metal ions. Furthermore, solvation and desolvation contributions, and the mutual spatial complementarity in van der Waals interactions are of utmost importance.[199, 201–203] The latter aspect has been summarized by Dunitz and Gavezotti in the context of attractive or repulsive interactions in molecular crystals as "empty space is wasted space".[204] Similar considerations have been found for protein–ligand complexes.[93] Additional effects are determined by intramolecular changes of receptor (R → R′ and ligand (L → L′) during complex formation.[194, 195]

2.2.1. Electrostatic Interactions

Pauling already highlighted the importance of hydrogen bonding[205, 206] for the structures of proteins and their ligand complexes.[207] Nevertheless, even today, no consensus view on the relative contribution of hydrogen bonding to the thermodynamics of protein folding and ligand binding has been established.[208–211]

Hydrogen bonds result from an electrostatic attraction between a hydrogen atom bound to an electronegative atom X (usually N or O) and an additional electronegative atom Y or a π-electron system. Distances of 2.5–3.2 Å between hydrogen-bond donor X and acceptor Y and X–H···Y angles of 130–180° are typically found.[206] Whereas no or only a slight dependency of the hydrogen-bond strength with angular changes are observed in the range of $180 \pm 30^\circ$,[212] shorter distances down to 2.3 Å result in a more covalent bond character and a larger binding energy,[213] although, the latter aspect does not hold in general.[214, 215]

As a result of the electrostatic nature, the strength of a hydrogen bond depends on its microscopic environment: the shielding of electrostatic interactions depends on the local dielectric constant ε of the surrounding medium, which means that the Coulombic interaction energy is proportional to ε^{-1}. While ε values of 1–20 (mostly 2–8) are assumed for the protein interior, the value at the protein periphery, next to the surrounding water, is about 80.[216, 217] Furthermore, in close proximity to polar groups a higher dielectric constant is expected compared to a nonpolar environment.[218] This also applies to more conformationally flexible regions of the protein.[219] Therefore, buried hydrogen bonds are regarded as more important for protein–ligand interactions than those formed in solvent-exposed regions.[220, 221]

The importance of water in the overall inventory of interactions is further indicated by the fact that only 1–2% of all buried N–H and C=O groups of protein amide bonds do not form a hydrogen bond.[222] Prior to complex formation, the functional groups of the uncomplexed receptor and the free ligand are involved in hydrogen bonds to surrounding water molecules in the solvent. In the complex, they are replaced by hydrogen bonds of comparable strength formed between the ligand and the receptor. It is hence the *difference* in the free enthalpies of these contributions to the hydrogen-bond

inventory that ultimately determines whether hydrogen bonding contributes to binding affinity or not.[223–225] Buried polar groups of a ligand or protein that remain unpaired are thus regarded as highly detrimental to complex formation.[220] As an upper limit in such unfavorable situations, a free-enthalpy contribution of 29 kJ mol^{-1} has been estimated.[226] On the other hand, these considerations underline why electrostatic interactions and hydrogen bonds are frequently the predominant contribution to the *specificity* of molecular recognition.[201, 210]

At physiological pH values (ca. 7.4), it is assumed that in proteins the guanidine side chain of arginine ($pK_a = 12.5$) and the terminal amino group of lysine ($pK_a = 10.8$) are protonated, whereas the carboxy groups of aspartic ($pK_a = 3.9$) and glutamic acid ($pK_a = 4.1$) are deprotonated (pK_a values according to reference [227]). Even more complex to predict are the properties of histidine residues ($pK_a = 6.5$). Their exact protonation state will depend upon the dielectric conditions imposed by the local environment. These can change upon ligand binding (Figure 3).[228, 229] If the bound ligand provides a sterically suitable arrangement of groups oppositely charged to the protein residues, attractive electrostatic interactions, so-called "salt bridges", can be formed (Figure 4).[230]

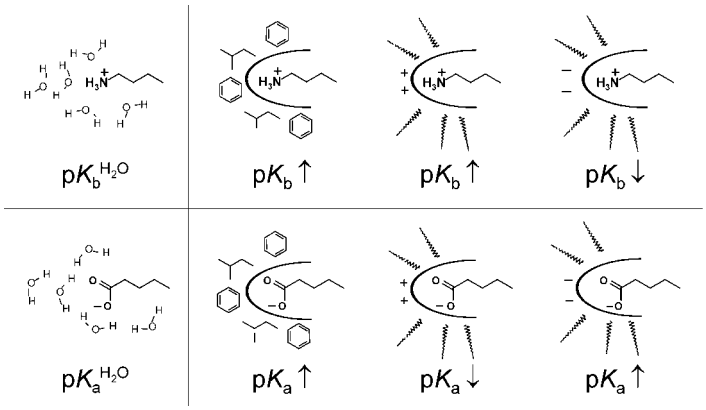


Figure 3. Impact of the protein environment on the pK_b values of a basic ligand group (upper row) and pK_a values of an acidic ligand group (lower row) compared to aqueous solution ($pK_b^{H_2O}$ and $pK_a^{H_2O}$).

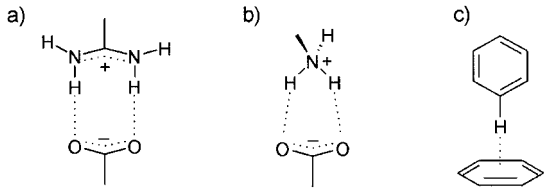


Figure 4. Examples of patterns of special hydrogen bonds: a), b) bidentate ionic ("salt bridges"), c) C–H $\cdots$ π interaction.

The contributions of hydrogen bonds and salt bridges to binding affinity have been estimated, however, the derived values must be considered with respect to the origin of these data. Evidence from protein-mutation studies suggest values for the interaction between uncharged partners of $\Delta G^\circ = -5 \pm 2.5$ kJ mol^{-1}.[210, 223, 231] Similar values originate from investigations of structures and solution energies of crystalline cyclic dipeptides[232] and studies estimating the contribution of intramolecular hydrogen bonds to the stability of proteins.[233–235] In contrast, values of -10 – -20 kJ mol^{-1} have been reported for charge-assisted hydrogen bonds and salt bridges.[231, 236] The interpretation of the experimentally determined "apparent" binding contributions suffers from one important problem common to all these studies: the measured quantities correspond to the "intrinsic" contribution of an interaction only if superimposed effects can be excluded.[231, 237, 238] For example, the contribution of a hydrogen bond was initially reported by Williams and co-workers to be about -25 kJ mol^{-1}.[239, 240] Later, this estimate was reduced to -1 to -7 kJ/mol^{-1} because of incorporation of initially neglected effects.[241, 242] Similarly, Andrews et al. overestimated the contribution of hydrogen bonds to complex formation because they assumed too large values for the entropic contributions that are detrimental to binding.[203]

Hydrogen bonds also influence ligand binding by their strong directional nature. Besides theoretical[243] and spectroscopic investigations, the analysis of crystal data primarily provides important information about their geometry.[244, 245] Carbonyl and carboxylate oxygen atoms form interactions mainly along the direction of their lone-pair electrons;[246, 247] for carboxylates the lone-pairs in the *syn* position are preferred to those with *anti* orientation.[244, 248, 249] A comprehensive compilation of the geometries of nonbonding interactions observed in crystalline solids is given in the IsoStar database[250] (Figure 5).

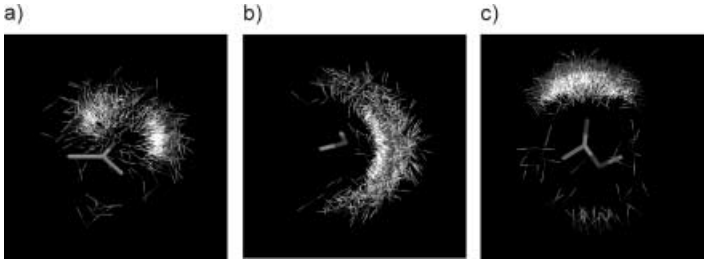


Figure 5. Composite picture of intermolecular interactions as observed in the crystal packing of small molecules compiled in the database IsoStar.[250] Shown are the arrangements of hydroxy groups around aliphatic ketones (a), aliphatic ethers (b), and aliphatic esters (c) as central groups.

Much weaker, but equally directional[251] hydrogen bonds are known between C–H $\cdots$ O, C–H $\cdots$ N, C–H $\cdots$ π-systems, and C–H $\cdots$ Cl[252–256] that also occur in more hydrophobic regions of proteins. Significant contributions to ligand-binding affinity also arise from so-called $\pi-\pi$ interactions[257, 258] between aromatic groups of ligands and side chains such as phenylalanine, tyrosine, or tryptophan.[259–261] Furthermore, interactions between cations, such as tetraalkylated amines, and aromatic residues are observed.[262–264] The latter play a significant role in the binding of positively charged ligands to the nicotinic acetylcholine receptor.[265, 266] Interestingly, comparative calculations on the strength of salt bridges versus cation–π systems in aqueous solution revealed an up to 10 kJ mol^{-1} greater contribution to binding affinity by the latter type of interactions.[267] Coordinative bonds of ligand functional groups (e.g. hydroxamates, carboxylates, phosphates, thiols) to protein-bound metal ions also stabilize

receptor–ligand complexes.[268, 269]. "More subtle" electrostatic contributions such as dipole–induced dipole, dipole–quadrupole, and quadrupole–quadrupole interactions to protein–ligand binding were found in highly resolved carbonic anhydrase II–ligand complexes. There, the electrostatic interactions were modified systematically by replacing the benzyl hydrogen atoms with fluorine atoms in *N*-(4-sulfamylbenzoyl)benzylamine.[270]

2.2.2. Solvation and Desolvation

In biological systems, molecular recognition between two molecules takes place in an aqueous environment. Thus, in addition to its role in the energetics of hydrogen bonds, as described in the previous section, water has an additional influence on the formation of protein–ligand complexes.[271–273]

In the condensed bulk phase, water molecules form a network of three to four hydrogen bonds per molecule.[205] Similar behavior is also found for about 80% of the water molecules that mediate interactions between protein and ligand, as evidenced by the analysis of 19 highly resolved crystal structures of protein–ligand complexes.[274] Assuming optimal geometry for these solvent-mediated interactions, a contribution to binding affinity of -10.5 to $-12.5\ kJ\,mol^{-1}$[275, 276] and $-7\ kJ\,mol^{-1}$[277] has been estimated. The latter value results from an entropic contribution ($-30\ J\,mol^{-1}\,K^{-1}$, corresponding to $9\ kJ\,mol^{-1}$ for $-T\Delta S$ at 298 K)[278] and an enthalpic contribution ($-16\ kJ\,mol^{-1}$)[279] which corresponds to the transfer of a water molecule from the bulk phase into the binding epitope.

Analyzing the topography of the surrounding molecular surfaces, it becomes evident that interstitial water molecules in the protein–ligand interface preferentially occupy cavities in the protein surface and less frequently reside in depressions on the ligand surface (Figure 6).[277, 280, 281] A series of highly

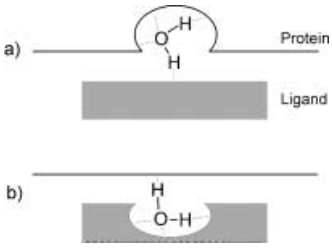


Figure 6. Schematic representation of water molecules mediating an interaction between protein and ligand; a) more frequently observed situation with a water molecule strongly bound to the protein; b) water molecule that is more strongly bound to the ligand (Figure adapted from reference [277]).

resolved crystal structures of the oligopeptide-binding protein (OppA) accommodating different Lys-xxx-Lys ligands (xxx: natural and non-natural amino acids)[282, 283] shows that, in these complexes, water molecules act as mediators to complement the side-chain residues xxx of different size, whereas the protein structure (apart from the rotation of the side chain in Glu32) remains almost unchanged (Figure 7).

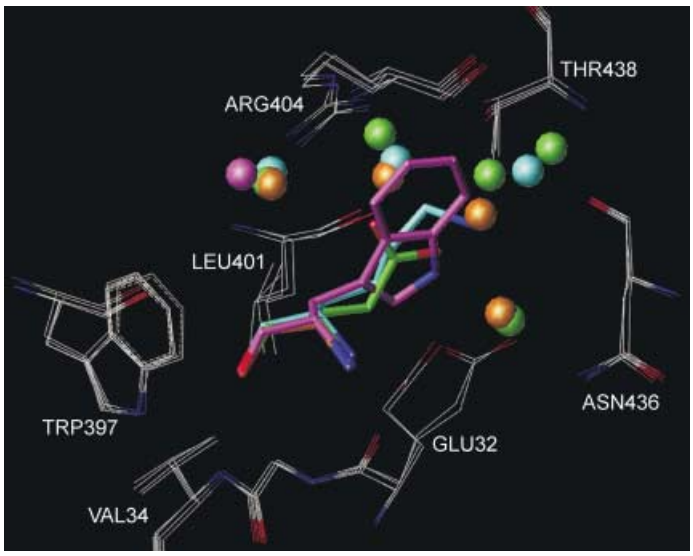


Figure 7. Superposition of the binding pockets extracted from four complex structures of the oligopeptide binding protein (OppA; white) with ligands Lys-xxx-Lys; for each example only the central xxx residue is shown in a different color (orange: Ala (PDB-Code 1jet), violet: Trp (PDB-Code 1jev), green: Glu (PDB-Code 1jeu), light blue: Lys (PDB-Code 2olb)). The water molecules are displayed as spheres with corresponding colors. The rigidity of the binding pocket (apart from Glu32) is obvious, together with a cluster formation of several water molecules.

However, the water molecules do not move without restriction within the ligand-binding pocket. In each case, they occupy energetically favorable, partially conserved positions (Figure 7).[277, 284–286] Interestingly, the binding constants vary in total by only one order of magnitude even for the exchange Trp $\rightarrow$Ala and Lys $\rightarrow$Glu as xxx.[283]

The unique role of water compared to other solvents—forming a tetrahedrally connected network and simultaneously occupying only an exceptionally small molecular volume[287–289]—also emerges in the desolvation of proteins and ligands upon complex formation. This step involves not only the rupture and reformation of hydrogen bonds to functional groups, but also the reorganization of the water structure at the interface. This is reflected both in the enthalpic and entropic contributions to the binding affinity.[240, 290–293]

The fact that the transfer of a nonpolar compound or a nonpolar surface portion into water is a) highly unfavorable, b) associated with a reduction in entropy at room temperature, and c) correlated to an increase in heat capacity has been summarized as the "hydrophobic effect".[233, 294–298] First introduced by Frank and Evans[299, 300] the "iceberg model" assumes that during the hydration of a nonpolar compound a *reduction* in the number of hydrogen bonds between water molecules occurs, but that water molecules next to the interface form *stronger* hydrogen bonds than those in the bulk water phase. Accordingly, Silverstein et al.[301] calculated a free enthalpy of $\Delta G = 2.0\ kJ\,mol^{-1}$ for the cleavage of a hydrogen bond in pure water, whereas the same step involving water molecules in the first solvation sphere around argon atoms requires a free enthalpy of $\Delta G = 2.6\ kJ\,mol^{-1}$. This effect results in a clathrate-type restructuring of the adjacent

water shell along with a partial immobilization of the water molecules.[302–304]

Whereas the enthalpic contribution almost cancels out for this process at room temperature (fewer but stronger hydrogen bonds instead of many such bonds of medium strength), the entropy decreases because of a higher ordering of the water molecules.[301] This step is entropically disfavored, but only up to a critical temperature T_s which depends on the nature of the compound to be transferred. At T_s the entropic contribution to the transfer vanishes.[297, 302] On the opposite side, for $T < T_s$, the burial of a hydrophobic surface upon complex formation corresponds to a favorable entropy-driven process ($\Delta H \approx 0$, $\Delta S > 0$). This view is supported by spectroscopic studies of surface-specific vibrations of molecular arrangements at the CCl_4/H_2O or hydrocarbon/H_2O interface. Although these studies indicate rather weak hydrogen bonds among the water molecules at the phase interface, the molecules mutually orient because of interactions with the organic phase.[305]

This classical view, however, is not generally accepted.[298, 302] An alternative approach does not regard the structuring of the water molecules as the main reason for hydrophobic interactions. Instead, it involves a positive enthalpy resulting from the rupture of several hydrogen bonds in order to create a cavity in the water structure that subsequently accommodates the nonpolar compound.[306, 307] In agreement with this hypothesis, calorimetric measurements found that contributions of 25 – 100 % of the protein – ligand binding enthalpy originate from solvent reorganization.[308]

The contribution of hydrophobic interactions to the free enthalpy in protein folding or protein – ligand complex formation can be regarded as proportional to the size of the hydrophobic surface buried during these processes.[309–314] This allows quantitative characterization of the effects involved.[315, 316] Solubility studies of hydrocarbons in water revealed -0.10 to $-0.14\ kJ\,mol^{-1}\,Å^{-2}$ as a contribution to hydrophobic interactions.[309, 310, 317, 318] The correlation of the hydrophobic surface buried upon receptor – ligand binding with experimentally determined binding affinities revealed values of -0.11 to $-0.24\ kJ\,mol^{-1}\,Å^{-2}$ as contributions to the free enthalpy.[240, 319, 320] The burial of a methyl group ($\approx 25\ Å^2$) contributes -2.75 to $-6\ kJ\,mol^{-1}$, thus increasing the association constants at 298 K by a factor of 3 – 11. In contrast, however, values of -0.08 to $-0.64\ kJ\,mol^{-1}\,Å^{-2}$ were determined in mutation studies for the influence of hydrophobic interactions on the stability of proteins.[321–324] Again, most values obtained suggest larger contributions compared with those resulting from solubility studies. This finding can be explained by a cooperative effect (see Section 2.2.4).[325] Once such effects are neglected, the sole consideration of the size of the buried surface reveals the reported higher contributions per $Å^2$. Huang and Chandler have recently suggested that for small hydrophobic molecules a scaling with molecular volume is more appropriate, whereas for larger hydrophobic molecules scaling with the molecular surface reveals a better correlation.[326]

Hydrophobic interactions are also regarded as the main driving force for conformational changes of the receptor upon ligand binding. This induced fit can be viewed as a "collapse" of the receptor about the ligand.[199] As an extreme case, the binding of trifluoperazine (**5**) to Ca^{2+} calmodulin induces a conformational change of the protein from an extended to a compact form (Figure 8).[327] Crystal structures of 3,4,5-substituted piperidine derivatives (e.g. **6**) bound to renin show an induced adaptation of the binding pocket to accommodate the attached substituents (Scheme 2).[328] Similar cases have been described for protein – ligand complexes of aldose reductase[329] and glycogen phosphorylase.[330]

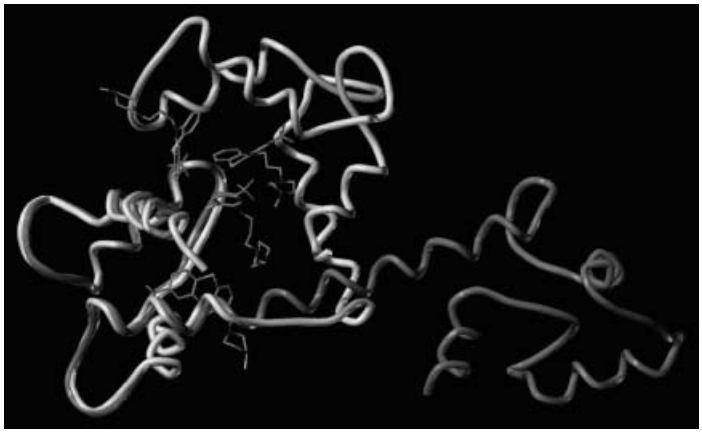
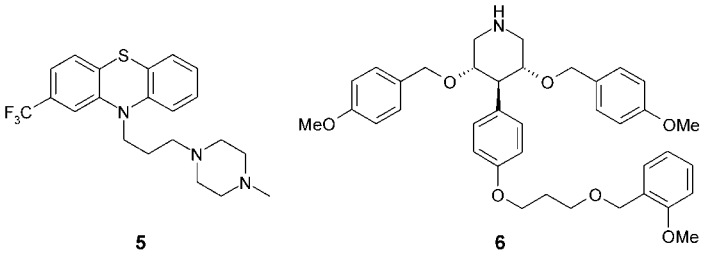


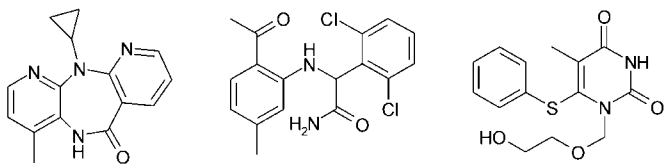
Figure 8. Superposition of the N-terminal domains of calmodulin as uncomplexed (dark gray: PDB code 1lin) and ligand-bound enzyme (light gray: PDB code 3cln). The four bound trifluoperazine molecules (**5**) indicate the induced collapse of the protein upon ligand binding. Only the backbone trace of the protein is shown in each case.



5 **6**

Scheme 2. Ligands of Ca^{2+}-calmodulin (**5**) and renin (**6**) that induce an adaptation of the protein binding pocket.

A comparable orientation of drugs with rather deviating shapes in the same binding pocket originates from induced-fit adaptations as a consequence of favorable hydrophobic interactions. An impressive example is the binding of nevaripin (**7**), α-APA (**8**), or HEPT (**9**; Scheme 3) to HIV-1 reverse transcriptase.[331] While none of the C-α atoms shift position by more than 2.7 Å, any structural adaptation of the protein is a response to changes in the substitution pattern of



7 **8** **9**

Scheme 3. Inhibitors of HIV-1 reverse transcriptase.

the inhibitors (Figure 9).[199] HIV-1 reverse transcriptase has to adapt its conformation consecutively while binding to the substrate RNA. Accordingly, it can be assumed that the observed inhibitor binding freezes different "snapshots" along this conformational transition path.

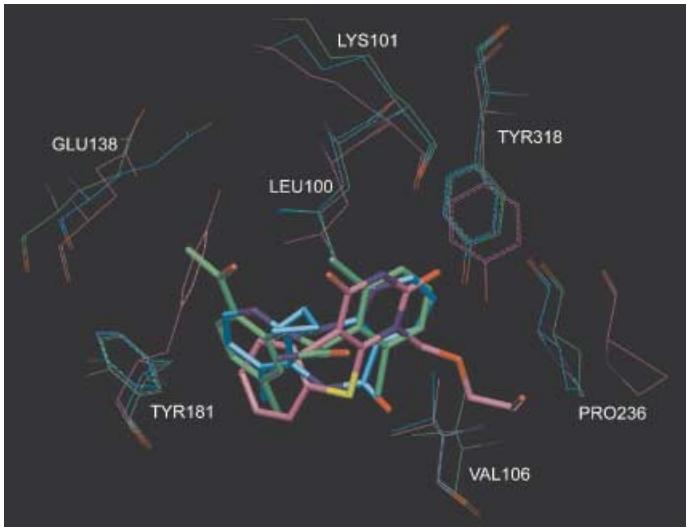


Figure 9. Induced fit of the binding pocket of HIV-1 reverse transcriptase resulting from the binding of nevirapin (**7**; cyan: PDB-Code 1vrt), α-APA (**8**; green: PDB-Code 1vru) and HEPT (**9**; violet: PDB-Code 1rti). The protein residues are color coded similar to the ligands. For clarity, only the amino acids which are next to the ligands in a sphere of < 4 Å and are involved in pronounced conformational changes upon binding of the different ligands are shown.

2.2.3. Intramolecular Changes of Ligand and Receptor

Upon complex formation, the change in degrees of freedom of the different components involved results in a change of entropy.[332] If the complex formation, neglecting any involved water molecules, is regarded as a bimolecular association step, each component loses three degrees of translation and rotation, while six new vibrational degrees of freedom are created.[290, 333–335] Although partitioning of the standard enthalpy into individual contributions is formally impossible for such processes in solution,[336] this simplification provides first insights into the influence of flexibility and mobility of protein and ligand on complex formation.[332]

The application of the Sackur–Tetrode approach or Trouton's Rule[337] requires the arguable assumption[338] that the results are transferable to solution processes[198, 339]. If so, for the *complete* immobilization of a molecule a loss in entropy of about $-420\ J\,mol^{-1}\,K^{-1}$ results.[340] However, losses about half this size ($\approx -200\ J\,mol^{-1}\,K^{-1}$[290, 333, 334, 343] corresponding to 60 kJ mol^{-1} at 298 K) were found experimentally. The difference can be attributed to the residual mobility of molecules in the complex. The latter is estimated by considering the experimentally observed motion in the crystals (e.g. lysozyme[341] or insulin[342]) or the entropy changes involved in inter- and intramolecular reactions. Even smaller contributions of 9–45 kJ mol^{-1} were found by Searle and Williams for the melting or sublimation of hydrocarbons or polar organic molecules.[292] This is in agreement with results for the association of rigid cyclic dipeptides in the solid, liquid, and gas phases.[344] Hermans and Wang calculated 29 kJ mol^{-1} (at 300 K) for the free enthalpy of binding of benzene to a lysozyme-T4 mutant resulting from the partial loss of translational and rotational degrees of freedom. This calculation also allowed the estimation of the remaining free space in the binding pocket available for motion of the benzene molecule. Positional deviations of 0.6 Å in atomic coordinates and 10–15° in rotation about the normal vector of the ring plane were found as root-mean-square deviations in a mutual superposition.[345]

Upon binding, conformational mobility is restricted, reducing the internal degrees of freedom of rotatable bonds. Such entropic contributions to the free enthalpy of binding were suggested to fall into a range between 0.5 kJ mol^{-1},[236] 2.5 kJ mol^{-1},[292, 343] and 4–6 kJ mol^{-1},[290, 333, 334] at 298 K. For amino acids, the probabilities of rotameric states were calculated based on observed conformational distributions of solvent-exposed side chains in protein crystals. These probabilities were used to estimate the entropy loss for the restriction of their conformational mobility.[346, 347] Contributions to the free enthalpy from 0 (for Ala, Gly, Pro) to 8.7 kJ mol^{-1} (for Gln), with a mean value of 3.7 kJ mol^{-1} per residue, were suggested.

Experimental evidence indicates that ligands frequently retain considerable residual mobility in the bound state. An *increase* in mobility of the protein can even favorably influence the free binding enthalpy.[348] Residual mobility is found, for example, for the binding of camphane, adamantane, or thiocamphor to cytochrome $P450_{cam}$.[349] Without forming a hydrogen bond, the ligands rotate freely in the binding pocket. They are, therefore, hydroxylated nonselectively. The binding of DNA to the C-terminal domain of topoisomerase 1[350] demonstrates that complex formation is not solely associated with a restriction of molecular mobility. While some of the protein residues become more highly ordered, others become more mobile. For the binding of a hydrophobic mouse pheromone to mouse major urinary protein, NMR relaxation studies[351] showed that an increase in protein backbone entropy reveals a considerable *favorable* contribution to the free enthalpy of binding, which is of the same order of magnitude as other contributions. Similarly, Weber et al.[134, 352] observed by crystallographic and thermodynamic studies of natural and synthetic streptavidin inhibitors that the ligand with the highest binding affinity also experiences the largest residual mobility in the complex.

An alternative strategy to compensate for entropic losses caused by ligand immobilization results from conformational preorganization in solution. In the case of the thrombin inhibitor D-Phe-Pro-boro-Arg,[353] a "hydrophobic collapse"[354] in aqueous solution minimizes the hydrophobic surface of the D-Phe and Pro side chains. At the same time, a conformation is adopted that strongly resembles the receptor-bound conformation.[355] Interestingly, an inverse "hydrophilic collapse" of the immune suppressors CsA and FK506 is made responsible for their high membrane permeability and suggests a formulation of the administered drugs in olive oil.[356]

Besides entropic contributions, enthalpic differences between solution and receptor-bound conformations of a ligand also influence the free binding enthalpy. Comparing computed force-field energies of the protein-bound conformations with those of the global minima in vacuum revealed differences between 0–167 kJ mol^{-1} for 33 compounds, with a mean value of 67 kJ mol^{-1}.[357] For three different dihydrofolate reductase inhibitors, an unfavorable conformational energy contribution to binding of 112–296 kJ mol^{-1} has been calculated.[358] These unrealistically high values result from a comparison of the receptor-bound conformers with those in the gas phase. However, such a comparison is of no relevance. Similar studies based on conformational ensembles produced from a solvation model suggest that conformational enthalpy differences amount to less than 12 kJ mol^{-1}, thus slightly disfavoring the receptor-bound state.[359] In addition, Vieth et al. found that the spatial orientation of "anchor points" responsible for ligand binding to the protein coincide well in protein-bound conformations with those of minimum-energy structures.[360] In several cases, conformations were also found for ligands which deviate slightly from those in the crystal structure, but have a significantly lower conformational energy.

It has to be considered that the anisotropic molecular environment of the protein perturbs the energy barriers separating different conformational (rotational) states. For example, this influence has been described by comparing the enzyme-bound conformation of methotrexate (an inhibitor of dihydrofolate reductase) with that adopted in its small-molecule crystal structure.[361] This polarization effect has to be considered in the development of advanced force fields.[362]

2.2.4. Additivity, Cooperativity, and Enthalpy–Entropy Compensation

Approaches based on group additivities [Eq. (4)] or the additivity of free enthalpy components [Eq. (5)] are frequently applied to understand and predict protein–ligand interactions. In this respect, pioneering work was performed by Andrews et al.[203] and Lau and Pettitt.[363]

$$\Delta G = \Delta G_{CH_3} + \Delta G_{OH} + \Delta G_{phenyl} + \ldots \quad (4)$$

$$\Delta G = \Delta G_{\text{H-bridge}} + \Delta G_{\text{solvation}} + \Delta G_{\text{conformation}} + \ldots \quad (5)$$

Already the variation in the absolute contributions discussed in the previous section demonstrates that this partitioning is not possible in a straightforward way. Strict application of statistical thermodynamics shows[364] that the free energy (free enthalpy) is a global property of the system under consideration. It thus depends on the total configuration space of the system. Hence, while separation of energy into (pairwise) individual contributions is a reasonable first approximation, this is, in principle, not valid for entropy[332] and free energy.[365] As a state function, the free energy is path independent, however, this does not apply for its components, as confirmed by nonadditivity in mutation studies.[237, 238, 366–368] A separation into individual components is possible if the total system under investigation is separated into mutually *independent* subsystems.[364] The latter is questionable, especially for biological systems featuring weak, noncovalent interactions which lead to many nearly identical (macroscopic) states.[369]

As an alternative, one can focus on the partitioning of the most dominant part of the free energy [Eq. (6)][365, 369, 370]:

$$\Delta G = \Delta H_{\text{H-bridge}} + \Delta H_{\text{solvation}} + \Delta H_{\text{conformation}} + \ldots + T\Delta S \quad (6)$$

This strategy has been used, for example, for the calculation of "intrinsic binding energies" from free enthalpies of binding of molecules with groups A, B, or A + B to a protein.[237, 291]

An impressive example for nonadditivity—also[371, 372] termed "cooperativity"—becomes apparent for the correlation of the "hydrophobic free enthalpy" with the solvent-inaccessible, nonpolar surface. Protein mutation studies and studies on ligand binding show that the hydrophobic effect obviously promotes stability and binding in aqueous solution to a large extent as suggested by solvent transfer measurements (see Section 2.2.2).[315] However, the burial of a part of a hydrophobic molecular surface at a binding site can induce a simultaneous cooperative enhancement of neighboring electrostatic interactions.[325, 372]

The contribution of the standard enthalpy ΔH° and entropy ΔS° to the free (binding) enthalpy ΔG° [Eq. (1)] can be determined directly from microcalorimetry[373] or van't Hoff plots of affinity measurements at different temperatures[374] (see Section 4). Generally, these experiments do not indicate any direct correlation between ΔH° and ΔG°. Thus, any interpretation or prediction of binding properties solely based on enthalpic considerations must be inadequate.[194] A possible exception might be given for series of closely related ligands with very similar entropic contributions.

The clear correlation between ΔH° and ΔS° ("enthalpy–entropy compensation") is obviously an intrinsic property of weak intermolecular interactions.[242, 336] This correlation is generally observed in (supramolecular) host–guest[375] and receptor–ligand complexes.[293, 376, 377] However, this form of compensation is by no means a "general" principle.[378, 379] It can be interpreted that an enhancement of intermolecular binding is accompanied by a loss in degrees of freedom of mobility and vice versa. Its existence is of particular importance for the prediction of receptor–ligand interactions: whereas the individual enthalpic and entropic contributions can vary over large ranges, the total change in free enthalpy is frequently close to zero. As a consequence, small *relative* errors in the prediction of ΔH° and ΔS° can have significant influence on ΔG°.

3. Theoretical Approaches to the Prediction of Binding Affinities

Studies on the prediction of binding affinities can be divided into two major categories:

- If the *3D structure of the biological target molecule is not known* the (often qualitative) prediction of the binding affinity of new ligands is based on the comparison with known reference structures such as endogenous ligands or

previously synthesized compounds.[150, 151, 380, 381] The considerable importance[382] of these methods arises from the fact that many pharmacologically important targets are membrane-bound proteins, such as G-protein-coupled receptors (GPCR),[383, 384] ion channels,[385] or transport proteins.[386–388] As, apart from a few examples, no experimentally determined 3D structure of sufficient accuracy is available for these systems, usually only indirectly obtained models can be used.

- If *the 3D structure of the receptor is known*, binding affinity predictions are performed considering geometrical and chemical complementarity between ligands and biological targets.[173–175, 198, 389–394] Because of the steadily growing number of spatially characterized receptors (Figure 2), an increasing impact of the latter strategy is to be expected in the near future.

In this review, we focus on methods exploiting the 3D structure of the receptor. As, however, comparative molecular-field analyses (3D QSAR) yield surprisingly good affinity predictions, merely by learning from the information provided by a ligand-training set, these methods will be briefly discussed in the following.

3.1. Approaches without Knowledge of the Receptor Structure

Affinity prediction of ligands in the absence of information about the receptor structure assumes that similarity in biological response is reflected by chemical similarity of the ligands.[380] Approaches that compare molecules on the one- or two-dimensional level[395] by means of topological descriptors consider the presence or absence of functional groups by associated bit vectors (so-called fingerprints). They will not be discussed here. Similarly, methods based on substructure mapping,[396] pharmacophore searches,[397] and superpositioning of ligands will not be considered.[398–400] Usually, these methods predict the expected binding affinity only on a qualitative scale.

In contrast, quantitative predictions can be obtained from Quantitative Structure–Activity Relationships (QSAR).[150] A correlation between structure and biological properties (e.g. affinity or selectivity) of a molecule is determined with respect to physicochemical and structural parameters. Classical 2D-QSAR methods, established by Hansch and Fujita[401] or Free and Wilson,[402] suffer from the fact that only data sets of structurally similar ligands can be studied and the spatial structure, essential for the understanding of receptor–ligand interactions, is only vaguely or indirectly considered.[403]

This limitation is relieved in 3D-QSAR methods:[152, 404] relative differences in the spatial structure of individual ligands are correlated with a known target property, such as binding affinity. As a prerequisite, the bioactive conformations of all ligands have to be considered to be aligned with each other, which best reflects their assumed arrangement in the binding pocket.[398] A conformationally rigid example in the data set could be used as a reference structure for a subsequent molecular superpositioning.[405] Equally well, conformations taken from known protein–ligand complexes,[406, 407] or functional groups in agreement with a pharmacophore hypothesis can be used.[408, 409] Besides atom- or group-based superposition methods, molecular fields are used to maximize their mutual similarity, in particular in flexible alignments.[398, 400]

In the following, a short description of current 3D-QSAR methods will be given, with emphasis on CoMFA (Comparative Molecular Field Analysis) and related developments. Comprehensive reviews in this field have been given by Kubinyi,[150, 151] Sanz et al.,[410] and van der Waterbeemd.[411]

- The 3D-grid-based CoMFA method,[405] developed from DYLOMMS,[412] compares a series of molecules in terms of molecular energy fields. It subsequently correlates field differences with differences in the dependent target property, for example, the binding affinity. In its original implementation, steric and electrostatic interaction energies are calculated for all molecules in the data set at the intersections of a grid containing all molecules. This approach assumes that entropic contributions are constant across the data set used for the analysis. For each molecule n, this QSAR is results in the Equation (7):

$$\text{Affinity}_n = k+\alpha_1 S_{n,1}+\ldots+\alpha_M S_{n,M}+\beta_1 E_{n,1}+\ldots+\beta_M E_{n,M} \quad (7)$$

The indices $1, 2, \ldots, M$ reflect the respective grid points, and $S_{n,1},\ldots, S_{n,M}$ and $E_{n,1}, \ldots, E_{n,M}$ describe steric and electrostatic energies at these points. The coefficients $\alpha_1, \ldots, \alpha_M$ and $\beta_1, \ldots, \beta_M$ are obtained from a system of linear equations by partial least-squares analysis.[413, 414] Binding affinities of new molecules, not included in the training set, can be predicted using the derived model.

While steric and electrostatic fields account for only enthalpic contributions, attempts have been described to reflect entropic contributions through the characterization of hydrophobic properties.[415] These approaches include fields based on HINT (Hydrophobic Interaction),[416, 417] molecular lipophilic potentials (MLP),[316] GRID fields[418] based on an H_2O or a DRY probe, or desolvation energy fields calculated with DelPhi.[419]

- Alternative molecular interaction fields are applied in CoMSIA (Comparative Molecular Similarity Indices Analysis).[420] Here, Gaussian functions are used to describe steric, electrostatic, and hydrophobic similarities. Similarly, hydrogen-bond donor and acceptor properties are considered.[244, 421] Compared to CoMFA, this approach avoids particularly steep potentials next to molecular surfaces. Thus, similarity indices are also determined close to the molecules.
- The HASL approach (Hypothetical Active-Site Lattice)[422] is another grid-based 3D-QSAR technique. It tries to attribute partial activities to grid points within the van der Waals volume of the ligands. The sum of the values at all grid points assigned to one molecule reflects the parameter to be correlated.
- In the Compass method,[423] molecular interaction fields are calculated in the proximity of the van der Waals surface of the considered molecules, thus focusing on the area likely to be involved in receptor–ligand binding. In addition, the number of descriptors is largely reduced. A QSAR model is

then generated using a back-propagation neural network. Subsequently, this model is further improved by iteratively generating and superimposing conformations of the molecules in the data set.

- In contrast to the previous methods, APEX-3D avoids molecular interaction fields.[424] Instead, a relationship between structural properties and observed activity is automatically established in a stepwise fashion. Molecules with similar activity are analyzed for corresponding 2D-topological or 3D-topographical patterns. Using logical programming, pharmacophores are identified which provide the basis for molecular superpositioning. Finally, a 3D-QSAR model is generated based on physicochemical properties of the pharmacophoric groups and global molecular properties such as hydrophobicity and molar refraction.
- The YAK method[158] is based on studies by Höltje and Kier.[425] Putative amino acid residues are placed in space using a set of ligands to generate a so-called pseudoreceptor. YAK selects and orients the amino acid side chains automatically. If available, crystallographic information, data from sequence analysis of homologous proteins, or mutation studies can be considered. This selection and positioning is iteratively optimized until the computed interaction energies best reflect the affinity data of the ligand under investigation. Finally, a pseudoreceptor is constructed by linking the placed amino acids with poly-Gly fragments.

3.2. Approaches Based on Knowledge of the Receptor Structure

The success of docking and de novo design methods strongly depends on the prediction of binding affinity, purely based on the spatial orientation of a ligand in the binding site.[174, 175, 426] The fundamentals of statistical thermodynamics used to calculate binding affinity are critically summarized in references [173, 198]. Reviews on applications are given in references [174, 175, 389–391, 394, 427, 428]. The determination of molecular interaction fields based on the known receptor structure is reviewed in reference [429]. References [430, 431] consider special applications and advances in force fields used in this context. A review on the handling of electrostatics in macromolecular systems is found in references [216, 432–436], whereas references [437–445] summarize the advances in calculating free enthalpy and entropy in the context of the thermodynamic perturbation theory.

In the following, the methods are classified and described with respect to their methodological background. This separation is not always strict, as several techniques combine different approaches.

3.2.1. Free-Energy-Perturbation Calculations and Linear Free-Energy Approaches

From a thermodynamic point of view, the rigorous prediction of relative free energies of binding of ligands to proteins results from free-energy-perturbation (FEP) calculations [Eq. (8)][446] or thermodynamic integration (TI) [Eq. (9)],[447] with explicit consideration of solvent molecules and flexibility of both the receptor and the ligand.[173, 390, 426]

$$\Delta F = F_1 - F_0 = -k_B T \ln\left\langle \exp\left(-\frac{H_1(\vec{X}) - H_0(\vec{X})}{k_B T}\right)\right\rangle_0 \quad (8)$$

$$\Delta F = F_1 - F_0 = \int_0^1 \left\langle \frac{\partial H_\lambda(\vec{X})}{\partial \lambda}\right\rangle_\lambda \mathrm{d}\lambda \quad (9)$$

The basis for these approaches is given by the relationship between the Helmholtz *free* energy F of a system and the ensemble average of an energy function describing the system under consideration.[444, 448, 449] $H_\lambda(\vec{X})$ is the energy of the system as a function of the coordinates $(\vec{X})$ of the particles in configuration space and a coupling parameter λ, k_B is the Boltzmann factor, and T the absolute temperature. The indices "0" and "1" represent $\lambda = 0$ and $\lambda = 1$, respectively. The configurational ensemble averages are taken either from Monte Carlo (MC)[450] or molecular-dynamics (MD) simulations.[451] As the difference between free *enthalpy* and free *energy* corresponding to the product of pressure and volume change experienced in an isothermal and isobaric reaction is negligible for processes in solution, free enthalpies are also available.

The method is suitable for studying individual contributions to the free energy/enthalpy on an atomic level or on the level of individual subsystems, such as ligand or protein.[238, 452, 453] However, it frequently encounters problems concerning the general applicability, which are caused by limited or insufficient sampling of the configuration space, the accuracy of the applied force fields, and the dependence of the results on the protocols used for simulation.[437, 438, 454] Moreover, long simulation runs are required and they can only allow for minor chemical differences in the ligands if their relative free energies/enthalpies are to be predicted reliably.[455–458] Some case studies, together with more recent approaches, will be mentioned.

- Postma et al.[448] and Jorgensen and Ravimohan[459] used FEP-MD and FEP-MC simulations to predict relative free-energy differences for the binding of benzamindine and *p*-fluorobenzamidine to trypsin,[460] or for a set of peptidic inhibitors to thermolysin.[225] In the latter case, a remarkable agreement between *predicted* and experimental values has been achieved,[461] although only minor structural modifications (exchange of NH to O by CH_2) of the ligands were studied. For example, the addition of a phenyl ring to an inhibitor obviously did not result in full convergence of the computed energies, even after 400 ps simulation time. Compared to experiment, the predicted relative free energy had the wrong sign.[457] Moreover, Graffner-Nordberg et al. stress that *all* processes combined with ligand binding—such as a change in protonation state—must be considered in the computed energies.[229] A comprehensive review of examples used for the prediction of protein–ligand affinities is given in reference [442].
- Ota and Brunger[462] combined non-Boltzmann sampling of configuration space with TI (NBTI). The advantage of this so-called umbrella sampling results from an artificially enhanced ligand flexibility caused by reduced internal

barriers to rotation and a thus augmented sampling of configuration space. Compared to classical TI, smaller deviations between calculated and experimental relative free energies are obtained for benzamidine or benzylamine binding to trypsin.[463]

- While standard free-energy calculations require time-consuming sampling of configuration space for *each* ligand modification individually, Gerber et al. suggest a simultaneous consideration of an entire *set* of modifications in one single MD simulation.[464] Assuming a linear separation of individual contributions, the derivative of each individual interaction with respect to the coupling parameter λ is determined analytically. Thus the initial gradients in free-energy contributions at $\lambda = 0$ allow the estimation of the contributions in the final state $\lambda = 1$. Although this method reduces computational efforts by a factor of 10^{-3}, simulations of the binding of trimethoprim-based inhibitors to dihydrofolate reductase and NADPH did not reveal a significant correlation between computed and experimental results.
- Oostenbrink et al.[465] used a "single-step perturbation" method to estimate relative free binding energies of endogenous and xenobiotic ligands to the estrogen receptor.[466] Instead of the usual FEP or TI calculations being performed for each ligand, an artificial reference molecule is simulated to generate a configuration ensemble that is representative of all ligands under consideration. Using Equations (8) and (9), the relative free enthalpy between two (real) ligands is calculated using the ΔG difference of the ligands with respect to the artificial reference molecule. Although the computational effort is reduced by a factor of 4–6 with respect to classical TI and mean deviations from experiment of 1.7 kJ mol^{-1} are obtained, it has to be considered that, for four out of five cases, ligand structural variations were limited to the presence or absence of hydroxy or methyl groups.
- Guo et al.[467] introduced a method in which the coupling parameter λ is handled as a dynamic variable. It develops together with the atomic coordinates of the system following Newton's law of motion. For a series of related ligands using a set of λs, relative free energies of binding are simultaneously computed. In the simulation, the different portions of the ligands all interact with the surrounding protein, but none of the ligands "takes notice" of any other one. An efficient mapping of configuration space is achieved with a significant reduction in computational effort.[468]
- To circumvent the problems occurring in classical free-energy simulations by using large, structurally diverse sets of ligands, a semiempirical method was developed by Åquist et al.[469, 470] They calculated absolute free energies of binding by considering MD simulations of two physical states. Polar and nonpolar contributions to the free energy are approximated linearly, by taking mean values from simulations of the ligand and protein–ligand complex in water. Required weighting parameters are calibrated using binding affinities of known complexes. However, this adjustment of energy contributions and the scaling of the weighting parameters depends on the simulation conditions and the system setup,[471, 472] and, accordingly, the general scope of the method appears limited.[473]
- Jorgensen and co-workers go one step further, by using an equation of the form (10):[474, 475]

$$\Delta G = \sum_i c_i \xi_i + \text{const} \qquad (10)$$

The physicochemical parameters ξ_i reflect ensemble average values obtained from MD or MC simulations. The parameters comprise, for example, the number of hydrogen bonds formed or the size of the buried hydrophobic, hydrophilic, and aromatic surface patches. The c_i values are adjusted by multiple linear regression using a training data set. In this respect the method can be classified as a regression-based approach (see Section 3.2.3). However, for the study of HEPT and nevirapin analogues binding to HIV-1 reverse transcriptase, differently composed Equations (10) were found, depending on the compilation of the training set.[475] This raises some suspicion about the general applicability and transferability of the approach.

3.2.2. Force-Field-Based Methods and Approaches Based on Additive Free-Enthalpy Contributions

The approaches described in this section assume partitioning of the free enthalpy of ligand-to-receptor binding into a sum of individual contributions [Eq. (5)] (for this assumption, see Section 2.2.4).[173, 198, 426] Starting from a "master equation" (ME), individual terms are defined on physicochemical grounds, whilst avoiding any cross correlations among them. Furthermore, unlike the methods described in the previous section (Section 3.2.1), all free-energy contributions are no longer derived as *ensemble* mean values, but taken from a single (or a few) generic structure(s). This is an important limiting approximation.[173]

- Modeling intermolecular interactions of protein–ligand complexes by simple molecular-mechanics force-field calculations in vacuum reflects purely enthalpic contributions to the free enthalpy of binding.[194, 426] By considering only van der Waals and electrostatic interactions, along with some intramolecular energy contributions, correlations with experimentally determined binding affinities have been obtained for series of closely related ligands, in which entropic contributions can be assumed to be constant.[476–480] In one example, the result obtained without the explicit consideration of solvent was explained by the predominance of van der Waals interactions and solvent-independent electrostatic contributions.[481]
- A straightforward approach to including solvent effects in the "master equation" which describes ligand–receptor binding is the use of atom-based solvation parameters,[311, 312, 314] usually scaled to the surface portion of protein and ligand that is buried upon complex formation. The methods of Vajda et al.,[482] Weng et al.,[483] Williams and co-workers,[240, 241, 484] Krystek et al.,[485] and Novotny et al.[486] consider contributions that are adverse to binding as additional terms, which originate from the loss of translational, rotational, and torsional degrees of freedom of the molecules (see Section 2.2.3). Krystek et al., Vajda

et al., and Weng et al. modeled the intermolecular interactions by Coulombic interactions using a distance-dependent dielectric constant; Williams and co-workers used intrinsic binding contributions of functional groups instead. For flexible ligands, Vajda et al. additionally[482] determined the energy difference experienced by *intra*molecular interactions of the molecule in the free and the bound state.

- The contribution of electrostatic interactions in the presence of water can be determined as an "averaged-field" approximation or continuum representation of the solvent by solving the linearized Poisson–Boltzmann equation[216] with the method of finite differences[487] or finite elements.[488] In this context, *polar* interaction energies of receptor, ligand, and receptor–ligand complex are compared to each other and determined with respect to the surrounding solvent, by considering the molecules (with discrete atomic charges) as regions of low dielectricity embedded in a medium of higher dielectric constant.[432] The *nonpolar* contribution to desolvation is assumed to be proportional to the size of the surface of both molecules buried upon complex formation. Entropic contributions attributed to the loss of mobility and flexibility are modeled as described in the previous section. Based on this concept, methods to predict free enthalpy of binding were suggested by Froloff et al.,[489] Zhang and Koshland,[490] Hofmann et al.,[491] Polticelli et al.,[492] and Shoichet et al.[493] Instead of the Poisson–Boltzmann approach, Zou et al.[494] used the "generalized Born Model" (GB/SA) of Still et al.[495] for the calculation of polar interaction energies.
- In the MM/PBSA method[440, 496, 497] and related approaches,[498–500] the free enthalpies for a molecular species are given by Equation (11):

$$\langle G \rangle = \langle E_{MM} \rangle + \langle G_{PBSA/GBSA} \rangle - TS_{MM} \quad (11)$$

E_{MM} reflects the mean molecular-mechanical energy, $G_{PBSA/GBSA}$ is the free enthalpy of (de)solvation obtained by solving the Poisson–Boltzmann equation (PB) or using the generalized Born approach (GB) and a surface-dependent term (SA). Both contributions are obtained by averaging over a sample of representative geometries extracted from a MD trajectory of the species under investigation with explicit consideration of water molecules and counter ions. The term $-TS_{MM}$ stands for the entropic contributions of the considered species, taken from a quasi-harmonic or normal mode analysis of the MD trajectory. The large scope of applications of this method demonstrates[501–504] that the properties of protein–ligand complexes exhibiting extensive structural differences can be studied.

- Alternatively, an implicit consideration of the contributions from solvation and desolvation can also be determined directly by molecular mechanics.[505] Thereby, the free solvation energy attributed to a functional group or amino acid residue is determined using the free solvation energy of the same group when part of a small molecule. This latter solvation contribution has to be reduced by an amount related to the exclusion of solvent molecules caused by their replacement by other atoms of the macromolecular system.

3.2.3. Regression-Based Approaches

As described in the previous section, regression-based approaches—also called "empirical scoring functions"—assume an additivity of individual terms to the total free enthalpy. However, the individual contributions (weighting factors or coefficients) of the separate terms describing the independent variables in the regression equation are determined either by multiple linear regression, partial least-squares regression [414] or a neural-network analysis,[506] using a training set of crystallographically resolved receptor–ligand complexes, together with experimentally determined binding affinities. Based on empirical concepts, the explanation of the obtained contributions along with their ability to predict unknown binding affinities justifies the initially assumed partitioning of the free enthalpy. Common to all regression-based methods, the results obtained as well as their transferability to new compound classes depend considerably on the compilation of the training set.[173] Furthermore, contributions of phenomena rarely observed in the experimental data, which frequently enough are the unfavorable ones, will be described insufficiently by the regression analysis.

- The archetype of an empirical scoring function for protein–ligand interactions was developed by Böhm (SCORE1).[507] Using a training data set of 45 protein–ligand complexes, the regression analysis with respect to experimentally determined affinities results in a cross-validated standard deviation of 9.3 kJ mol^{-1}. In this analysis, the sum of contributions to hydrogen bonds, ionic interactions, buried nonpolar surface regions, and the loss of (intra)molecular mobility has been considered. Expanding this training set to 82 complexes and considering the degree of burial of hydrogen bonds in the binding epitope along with special terms for aromatic and unfavorable electrostatic interactions, a standard deviation of 8.8 kJ mol^{-1} has been achieved for a prediction data set (SCORE2).[508] These analyses also showed that the (relative) contributions of the individual terms depend on the compilation of the training set. The same holds for different strategies in partitioning the free enthalpy.

  A similar approach has been described by Eldridge et al.[509] (82 complexes in the training set, ChemScore) and Wang et al.[510] (170 complexes in the training set, SCORE). Compared to Böhm's approach,[507, 508] Eldridge et al.[509] handle contributions for intramolecular flexibility differently and Wang et al.[510] classify hydrogen bonds as "strong, moderate, and weak", including the occurrence of interstitial water molecules as mediators of interactions. An "evolutionary test" demonstrates[510] that the resulting coefficients only converge if a set of more than 100–120 protein–ligand complexes, which deviate sufficiently in the types of intermolecular interactions, is used for analysis. Murray et al.[511] improved the predictive power of the scoring function obtained by Eldridge et al.[509] with respect to a *selected* protein including additional information through Bayesian statistics.
- In their "VALIDATE" approach, Head et al.[512] use electrostatic and steric interaction energies from AMBER,[513] an HlogP-based[514] octanol–water partition coefficient, polar

and nonpolar contact surfaces, and a term to describe intramolecular flexibility. The coefficients for the various contributions were derived based on 55 protein–ligand complexes by means of a partial least-squares[414] or a neuronal-net analysis. Both strategies result in regression equations that can hardly be interpreted in physical terms. Their relevance, in particular with respect to ligand *optimization,* is thus rather limited.

- In studies of Takamatsu and Itai[515] (29 avidine ligand complexes in the training set), Venkatarangan and Hopfinger[516] (23 glycogen phosphorylase–inhibitor complexes in the training set), and Viswanadham et al.[517] (11 HIV-1 protease–inhibitor complexes in the training set), AMBER interaction energies between protein and ligand[513] were combined with additional terms to describe hydrophobic interactions and other entropic contributions. The individual coefficients were again determined by multiple linear regression.
- Rognan et al.,[518] Bohacek and McMartin,[519] and Kasper et al.[520] also developed empirical scoring functions tailored towards *one* particular protein. In the first case, training sets of five crystallographically determined HLA-A*0201 peptide–inhibitor complexes and 37 modeled H-2K^k peptide–inhibitor complexes were used to reparameterize the scoring function of Eldridge et al.[509] Bohacek and McMartin used only nine characterized thermolysin–inhibitor complexes for calibration and considered only the number of hydrogen bonds or hydrophobic contacts in their scoring function. Kasper et al. used an approach similar to Equation (11) scaling, however, the different contributions relative to a training set of 11 peptide–chaperone DnaK complexes.
- In contrast to the functions described above, Jain developed a function[521] that is continuously differentiable. Terms to describe hydrophobic and polar complementarity of receptor and ligand are modeled combining a Gaussian and a sigmoidal function. Only ligand-dependent contributions are used for handling entropic considerations. The analysis is based on 34 protein–ligand complexes in the training set.

3.2.4. Knowledge-Based Approaches

Knowledge-based approaches are based on the idea that a sufficiently large data sample can serve to derive rules and general principles inherently stored in this knowledge base.[522] Accordingly, the development of a knowledge-based scoring function at an atomic level is based upon observed frequency distributions of typical interactions in experimentally determined structures: in any system, only those interactions that are close to the frequency maxima of the interactions in the knowledge base are considered to be favorable. This approach has been successfully applied in the field of protein-fold prediction.[523–525] Using the concept of the "inverse Boltzmann law",[526] the frequency distributions of interatomic interactions, derived from protein crystal structures, are converted into "potentials of mean force" or "knowledge-based potentials". Although the thermodynamic foundation of this procedure[365, 527–529] and the terminology used[530] have been debated, the results obtained by these approaches are superior to those obtained by molecular-mechanics force fields.[531–534]

Recently, the following approaches have been published on protein–ligand systems:

- Verkhivker et al.[535] derived distance-dependent knowledge-based pair potentials for a data set of 30 HIV-1, HIV-2, and SIV protease–inhibitor complexes. These potentials were combined with desolvation terms for ligand and protein using atom-based parameters.[536] To estimate contributions arising from the conformational immobilization of protein side chains, a method introduced by Pickett and Sternberg has been adapted.[346] Differences in binding affinities of several HIV-1 protease–inhibitor complexes can be reproduced by this concept.
- Surface patches of *pairs* of interacting atoms buried upon complex formation are computed by Wallqvist et al.[537] in terms of frequency distributions from a set of 38 protein–ligand crystal structures. Atom-based statistical preferences are produced by normalizing with the product of buried surfaces of the corresponding *individual* atoms. Using two parameters calibrated by experimental binding affinities of the training-set molecules, binding affinities of ten additional HIV protease–inhibitor complexes were predicted with a standard deviation of 6.3 kJ mol^{-1}.
- DeWitte and Shakhnovich[538] used 17 or 109 crystal structures, respectively, from the protein data bank (PDB)[2] to develop "interatomic-interaction free energies" (SMoG-Score) for ligands that bind to the surface of a protein or into binding pockets. Using a Metropolis–Monte Carlo-based[450] construction procedure, ligands are generated and energetically ranked in the binding pocket. The method has been applied to complexes of purine nucleoside phosphorylase, the SH3 domain, and HIV-1 protease.
- Muegge and Martin[539] produced "Helmholtz free interaction energies" ("PMF" score; potential of mean force) from 697 crystallographically determined protein–ligand complexes by using 16 protein and 34 ligand atom types, respectively. Implicit contributions of water are considered using a specific volume correction term[540] and sampling atom distances up to 12 Å to produce the pair-distribution functions. For a test set of 77 protein–ligand complexes studied crystallographically, a deviation of 1.8 log units in reproducing the experimentally determined binding constants is found.
- Mitchell et al.[541] published pair potentials (BLEEP) derived from 820 protein–ligand atom-pair distributions based on the "inverse Boltzmann approach". The analysis included hydrogens initially positioned by the program HBPlus.[222] As reference state, a semi-empirical Ne–Ne pair potential, suggested by Ng et al.[542], has been used. In addition, the consideration of water molecules as part of the protein has been tested. For 90 diverse protein–ligand complexes a correlation coefficient of 0.74 (a standard deviation is not reported) is achieved for experimentally determined affinities.[543]
- The scoring function DrugScore, developed by us,[544] (see Section 5) is composed of distance-dependent pair poten-

tials and solvent-accessible-surface-dependent singlet potentials. They are computed using distribution functions retrieved from 1376 protein–ligand complexes as stored in the ReliBase data base.[545] For 55 well-distributed protein–ligand complexes determined by crystal-structure analysis, a deviation of 1.8 log units from experimentally determined inhibition constants has been found.[546]

3.2.5. Consensus Scoring and Filter Functions

Although not satisfactory from a scientific point of view, a pragmatic strategy to enhance the reliability of predicted binding affinities results from the simultaneous consideration of several scoring functions. Charifson et al.[547] used a logical AND operation to combine the scoring of ChemScore,[509] the AMBER-based[513] function in DOCK,[548] and the "piecewise linear potential" function[549]. In a virtual screening assay using three different target enzymes, this consensus scoring function allowed the retrieval of known active inhibitors from a set of randomly selected molecules with a significantly improved reliability. So and Karplus averaged the predictions of up to five different QSAR methods and showed that the combined predictions were superior to the results obtained from the individual evaluations.[550] Using seven different target proteins for virtual screening, Stahl and Rarey[551] considered a combination of terms from PLP score[549] and SCORE1.[507] They implemented this function into the docking program FlexX[552] to achieve overall more robust enrichment rates. Interestingly enough, in several cases the combined function does not achieve the same (high) enrichment rates as obtained with the original functions. Terp et al.[553] even proceeded a step further by correlating the scores of eight functions by using a PLS analysis with experimentally determined binding affinities for a heterogeneous data set of 120 crystallographically determined protein–ligand complexes. Compared to the Consensus Scoring suggested by Charifson et al., the latter approach provides *quantitative* predictions of binding properties. Applying this model to predict affinities of 120 docked MMP inhibitors revealed deviations of less than one log unit for p*K*i values in 49% of the cases.

Because of the way in which they are derived (see Section 3.2.3) regression-based methods evaluate predominantly those favorable interactions *most frequently* exhibited in crystal structures of protein–ligand complexes. However, to recognize and disfavor those protein–ligand geometries occasionally produced by computational docking but which are not in agreement with the experimental evidence, Stahl and Böhm suggested the usage of "filter functions".[220] These functions cope, for example, with situations in which polar atoms are buried upon binding but do not form appropriate hydrogen bonds, or where hydrophobic cavities in the protein–ligand binding epitope are generated.

3.2.6. Approaches to the Location of Interactions in Space

Assuming a successful partitioning of binding affinity into individual (additive) contributions (see Section 2.2.4), methods for locating favorable interaction sites can play an important role in the optimization of ligands in the binding pocket.

- The archetypical method in this area is Goodford's GRID program.[418, 554] It is based on a tailored force field. Regions in the binding pocket are contoured in terms of the interaction energies that various probes experience at the intersections of a regularly spaced grid. Such probes could be, for example, water, amino or carboxy groups, or hydrophobic groups (DRY).
- Similar concepts, but based on crystal data are used in the X-SITE[555] and SuperStar[556, 557] methods. The X-SITE method uses spatial contact distributions derived from 163 triatomic fragments to highlight favorable interaction sites in a binding pocket. The distributions were retrieved from 83 high-resolution protein structures (without ligands). SuperStar uses spatial information stored in IsoStar.[250] This latter database comprises nonbonded interactions compiled from crystal data of small molecules in the CSD.[1] These data are subsequently used to calculate probability densities for contacts with atoms of functional groups (such as ammonium nitrogen atoms, carbonyl oxygen atoms, methyl carbon atoms, …) at the intersections of a grid embedded into the protein binding pocket.
- Similarly, the knowledge-based pair potentials implemented into DrugScore can be used to identify hot spots of binding[544] (see Section 5) using appropriate probe atoms (such as aliphatic carbon, carboxylate, carbonyl, hydroxy-oxygen, amino nitrogen…).[546]
- In contrast, the MCSS (Multiple-Copy Simultaneous-Search) approach[558] based on the CHARMM force field[559] distributes probe molecules such as acetamide, methanol, acetate, or propane at favorable positions in the binding pocket. The method has been extended to study flexible regions in the binding pocket.[560]
- The PROFEC (Pictorial Representation of Free-Energy Changes) approach of Radmer and Kollman[561] and enhancements of Pearlman[562] (OWFEG, One-Window Free-Energy Grid) are based on FEP calculations. Two MD trajectories are used to determine free-enthalpy changes resulting from the placement of an atom or group at *different* locations around an inhibitor, both in solution and at the protein binding site.

3.2.7. Comparison of the Various Approaches

A comparison of the developed methods with respect to quality and speed is difficult. First of all, there is not yet a generally accepted data set for establishing and testing a new method. The hardware requirements and the necessary data input preparation for the different approaches are difficult to compare. Moreover, frequently enough the authors have studied a limited set of examples with respect to the scope of the biological systems considered. Accordingly, a reliable assessment of the different methods is difficult. Despite these limitations, the published methods for affinity prediction summarized in Sections 3.2.1–3.2.4 are compared in Table 2 from a methodological point of view. In addition, relationships to other studies or methods are listed.

Table 2. Comparison of the methods for the prediction of binding affinity of receptor–ligand complexes with knowledge of the 3D receptor geometry discussed in Sections 3.2.1–3.2.4.

First author (method name)	Reference	Cross references[a]	Method[b]	Number of test systems[c]	SD[d]	Time	Comments
Wong	[460]	–	FEP-MD	2	2.2	–	replacement of benzamidine by *p*-fluorobenzamidine and mutation of Gly216Ala in trypsin
Reddy	[564]	–	FEP-MD	2	3.6	–	replacement of a formyl group with a proparagyl group
Bash	[225]	[461]	FEP-MD	2	0.5	–	replacement of a NH group with an O atom
McCarrick	[457]	–	FEP-MD	3	8.4	–	substitution of phenyl rings
Ota	[463]	[462]	NBTI	2	1.7	–	improved sampling of the configurational space; replacement of benzamidine with benzylamine
Gerber	[464]	–	derivations of the free energy	2×36	–	acceleration by 1000^{-1}	no significant correlation between experimental and calculated affinities
Oostenbrink	[465]	[466]	single-step FEP-MD	5	3.3[e]	acceleration by 5^{-1}	substitution of hydroxy and methyl groups in four of the five ligands.
Guo	[467]	[468]	λ-dynamic approach	4	2.1	–	replacement of benzamidine with *p*-aminobenzamidine, *p*-methylbenzamidine, *p*-chlorobenzamidine
Åquist	[469]	[471–473]	LIE	18	3.9	–	SD value of Model 6 in Table 2 from reference [470]
Rizzo	[475]	[474]	LIE/LR	2×20	3.9	–	different regression equations each according to composition of the data sets
Grootenhuis	[476]	[477, 479]	CHARMM energy	35	8.3	2–5 min per compound	SD value from Protocol 8, Table 3
Holloway	[478]	–	MM2X energy	15	5.7	–	SD for test set from Table 2
Vajda	[482]	[486]	ME based	$9+3+5+9$	5.4	–	SD for test set from Table 1
Wenig	[483]	[482]	ME based	$9+10+8$	≈ 4.2	–	investigation of protein–protein complexes
Williams	[484]	[240, 241]	ME based	1	≈ 11.4	–	SD estimated from errors of individual contributions
Krystek	[485]	–	ME based	9	16.7	–	SD estimated from errors of individual contributions
Checa	[481]	–	AMBER energy + PBE	7	3.3	–	AMBER energy alone correlated equally significantly
Froloff	[489]	–	PBE + ASP	$3+5$	>42	–	SD for test set from Table 2 in ref. [489]; systematic error
Zhang	[490]	–	PBE + ASP	9×7	2.1	–	9 mutants of isocitrate dehydrogenase as protein components
Kuhn	[502]	[440, 497]	PBE + ASP	1	16.7	–	SD for the biotin/avidin complex
Hoffmann	[491]	[565]	CHARMM + PBE + ASP	10	–	"several h for 100 compounds"	improvement in the placing of docked geometries as target
Polticelli	[492]	–	PBE + ASP	4	56.4	–	SD for test sets from Tables 1 and 2; systematic error
Shoichet	[493]	[548, 566]	Born equation + ASP	5	20.9	–	SD for test set in Table III; systematic error
Zou	[494]	[566]	GB/SA	6	6.3	10 s per compound	SD for parameter set 1 in Tables 2 and 3
Böhm (SCORE1)	[507]	[565, 567]	regression based	45	9.3	"several compounds per second"	cross validated SD for function 2
Böhm (SCORE2)	[508]	[565, 567]	regression based	$82+12$	8.8	"several compounds per second"	SD for test set from Table 3
Eldridge (ChemScore)	[509]	[511]	regression based	$82+20+10$	8.7	–	cross-validated SD for total training set from Table 8
Wang (SCORE)	[510]	[568]	regression based	170	6.3	–	cross-validated SD for total training set from Table 6
Head (VALIDATE)	[512]	–	regression based	$51+14+13+11$	6.3	–	cross-validated SD for total training set from Table 2
Takamatsu	[515]	–	regression based	29	–	–	calibration solely on avidin complexes
Hopfinger	[516]	–	regression based	15	–	–	calibration solely on glycogen phosphorylase complexes
Viswanadhan	[517]	–	regression based	11	2.4	–	calibration solely on HIV-1 protease complexes
Rognan	[518]	[509]	regression based	$5+37$	3.1 or 5.1	–	calibration on HLA–A*0201 and H-2–K^k complexes; SD given in each case.
Bohacek	[519]	–	regression based	9	2.3	–	calibration solely on thermolysin inhibitors
Kasper	[520]	–	regression based	11	1.7	–	calibration solely on DnaK–Heptapeptide complexes
Jain	[521]	–	regression based	34	5.7	–	cross-validated SD for function "F"
Verkhivker	[535]	–	knowledge based	7	–	–	derivation and test of the function solely on HIV and SIV proteases
Wallqvist	[537]	[569]	knowledge based	8	6.3	–	SD for calibration set in Table 3
DeWitte (SMoG-Score)	[538]	–	knowledge based	$17+8+11$	–	–	no SD value given
Muegge (PMF-Score)	[539]	[570, 571]	knowledge based	77	10.3	–	SD for test set 6 in Table 4
Mitchell (BLEEP)	[541]	[543, 572]	knowledge based	90	–	–	no SD value given
Gohlke (DrugScore)	[544]	[546]	knowledge based	71	9.2	0.2 s per compound	SD for test set "Böhm1998" in Table 2

[a] Listed are references to related work or applications of the method in other programs. [b] For detailed explanations of the methods listed see text. FEP–MD = free-energy perturbation/molecular dynamics, NBTI = non-Boltzmann thermodynamic integration, LIE = linear interaction energy, ME = master equation, PBE = Poisson–Boltzmann equation, ASP = atomic solvation parameter, GB/SA = generalized Born approach, LR = linear regression. [c] The number of protein–ligand complexes used for validation in the individual test sets are reported. [d] The standard deviation between calculated and experimental binding affinities are reported (in kJ mol^{-1}; a temperature of 298 K was assumed for conversion of binding affinities given in logarithmic units). [e] Given is the mean deviation for *relative* free binding energies.

Currently, only a few comparative studies on the evaluation of affinity prediction methods are available. In all published comparisons[547, 551, 563] enrichment rates achieved in virtual screening are used to assess the quality of the predictions but not the accuracy of the achieved affinity predictions. These studies conclude that, currently, no general-purpose function is available. Depending on the type of target protein and the predominating protein–ligand interactions, the best-suited method has to be selected based on some preliminary tests.

4. Experimental Approaches to Describing Binding Affinity

4.1. Indirect Methods

Binding affinities are usually determined in a binding assay. In the case of enzyme reactions, the influence on enzyme kinetics is followed by means of a readily detectable physical property (e.g. absorption, fluorescence, or fluorescence polarization of one of the reaction partners). The inhibition or binding constant of a ligand is subsequently derived indirectly, by considering the changes in concentration or changes in enzyme kinetics, respectively. For receptor binding studies, inhibitor binding is recorded by the replacement of the ligand with potent, suitably labeled compounds. In all cases, an indirect determination of the binding constant and, accordingly, the free binding enthalpy ΔG° is performed. The partitioning of standard enthalpy ΔH° and standard entropy ΔS° to ΔG° [Eq. (2)] can be determined using van't Hoff plots of affinity measurements at different temperatures.[573]

In recent years, a large variety of physicochemical methods has been established for the quantitative determination of protein–ligand binding; they are currently being developed further. In this review, only a few such methods will be discussed exemplarily. Plasmon resonance spectroscopy can detect binding of a ligand to a protein immobilized on a solid support or, vice versa, of a protein to a ligand attached to the support by an appropriate anchor group.[574–576] In particular, the "on" and "off" rates of binding can be studied using this method. A number of NMR spectroscopic pulse sequences have been developed to detect binding through signal shifts or by recording the transfer of magnetization between protein and ligand.[577–582] Mass spectrometry allows conclusions about the stability and binding affinity of protein–ligand complexes by determining complex dissociation as a function of the measurement parameters (e.g. acceleration voltage).[583] Under such conditions, the binding parameters are recorded by excluding the solvent environment. This provides valuable complementary information to the other methods.[584] Furthermore, atomic-force microscopy has been used to determine the strength of protein–ligand interactions by a controlled rupture of the studied complexes.[585]

4.2. Direct Measurement of Thermodynamic Parameters

Direct access to binding affinities is accomplished through microcalorimetric measurements.[373] Because of its importance for the understanding of the thermodynamics of ligand binding, this method will be discussed in more detail. In isothermal titration calorimetry (ITC), a ligand is added in a stepwise fashion at constant temperature to a buffered solution of the receptor. The overall heat of reaction generated upon complex formation is recorded.[586] The association constant K_A (and, accordingly, ΔG°), together with the stoichiometry of the ligand–receptor binding process are also available. In addition, the binding constant can be computed from the shape of the titration curve. The enthalpic portion of the binding process can be derived from the integrated heat of reaction, the entropic contribution $T\Delta S^\circ$ can be calculated from the difference between ΔH° and ΔG°. Reliable shape analysis of the titration curve requires binding constants of $< 10^9\,\mathrm{M}^{-1}$. To measure compounds of higher affinity, the detection range can be extended by the displacement of a lower affinity ligand.[587–589] Alternatively, for larger affinities the binding constants can also be taken from other experiments (e.g. enzyme kinetics).

The heat of reaction measured in an ITC experiment comprises all intermediate and transient reactions that are superimposed on the binding process. Measuring under different buffer conditions elucidates whether a proton transfer step is involved in the binding process.[586, 590–592] The functional groups of the protein or ligand can experience changes in protonation state upon complex formation. As a result, the transfer of a ligand from aqueous solution to the protein environment can strongly affect the dielectric properties of the local environment of these groups, resulting in a significant shift of pK_a (pK_b) values (Figure 3). The measured total enthalpy ΔH_{total} is thus composed of the reaction enthalpy ΔH_{bind} and the enthalpy attributed to the proton exchange reaction with the buffer medium ΔH_{ion}. As different buffers exhibit different ionization enthalpies,[593] this process can be readily detected by measuring under several buffer conditions [Eq. (12)][586]

$$\Delta H_{total} = \Delta H_{bind} + n\Delta H_{ion} \qquad (12)$$

The stoichiometry of the protonation reaction is available from n, with the sign of n indicating whether protonation or deprotonation of the ligand and protein groups is occurring.

From van't Hoff plots used to analyze affinity data measured at different temperatures, only that part of the binding enthalpy can be extracted that refers directly to the parameters which determine the observed measurement signal (e.g. absorption, fluorescence quenching). This means that only for a direct transformation of the system from a properly defined incipient state (ligand and protein separated) to a final state (protein–ligand complex) the extracted enthalpy is equivalent to the value of ΔH_{bind} obtained by ITC. No intermediate (e.g. conformational) states are allowed to occur, nor should other steps (e.g. change in the protonation state) be superimposed onto the binding process.[594] In biological systems, ΔG usually shows only a low temperature dependence (see below). Thus, a reliable determination of enthalpy and entropy is hardly possible from van't Hoff plots.

The interpretation of ITC results gives rise to a number of interesting aspects, especially with respect to structural details

of receptor–ligand binding.[197] In addition to the detection of superimposed protonation reactions, the confirmation of a virtually temperature-independent value of ΔG° as a result of a pronounced enthalpy–entropy compensation has to be emphasized (see Section 2.2.4).

Further access to structural interpretations is provided by the determination of heat-capacity changes ΔC_p. ITC measurements carried out at different temperatures show a strong temperature dependence of ΔH° and, conversely, of $T\Delta S^\circ$ for most biological systems. In such cases, analogous van't Hoff plots show a nonlinear behavior of $\Delta G^\circ/T$ as a function of the reciprocal temperature. In contrast to enthalpy and entropy changes, ΔC_p is virtually temperature independent in the range usually accessible to biological systems. Negative values are found in general for protein–ligand complexes. Accordingly, the complex exhibits lower heat capacity compared to the sum of the free components. With respect to enthalpy and entropy, this general behavior means that, with increasing temperature, protein–ligand binding becomes increasingly exothermic and simultaneously entropically less favorable. Any interpretation in terms of enthalpy- or entropy-driven binding must therefore be considered in the light of the applied temperature conditions. Several models have been discussed in the literature that correlate heat-capacity changes with the hydrophobic surface accessible to water molecules prior to binding but buried upon protein–protein complex formation or protein folding.[590, 595–600] Similar surface-dependent contributions were found for the transfer of hydrophobic solvent molecules from the water phase into their pure phase (see Section 2.2.2). If this empirical correlation model is applied to the binding of small ligands to their receptors, the values calculated for ΔC_p are too small.[601–603] Accordingly, the experimentally determined heat capacity of the complex formation is smaller than the value predicted in comparison to the separated components, if the model used only considers the surface contributions buried upon complex formation. To calculate such surface contributions, the relevant structure of the free and ligand-bound protein has to be known. In general, crystal structures are used for this purpose. Therefore, for the binding of a small-molecule ligand to its receptor, the surface-dependent release of the hydration shells cannot provide the sole contribution to the change in heat capacity. Conformational transitions of the binding partners or differences in the excitable vibrational modes of the macromolecular structures arising from complex formation have been discussed as additional explanations.[604–606] The deviations are, however, also observed for conformationally rigid proteins such as trypsin[607] and obviously represent a general phenomenon in protein–ligand binding. Thus, it appears likely that they involve water which is ubiquitously present in all binding processes. Liggins and Privalov[608] assume that the enthalpic contributions to binding resulting from hydrogen bonds formed in the complex are exaggerated. The enthalpy of dehydration necessary for complex formation is also included in this contribution. At the interface of a binding pocket and bulk solvent, the assumption of complete dehydration is probably an overestimate. More likely, even after protein binding, polar ligand groups still influence the water structure in the local proximity. This effect leads to reduced desolvation contributions which could possibly influence the heat-capacity changes. It might explain the discrepancies in the empirical relationships established for protein-folding experiments and the calculated and the measured ΔC_p values for ligand–protein complexes.

Further information on heat capacities and their changes can be obtained from DSC measurements (differential scanning calorimetry).[594] The system under investigation is heated under virtually adiabatic conditions at a constant rate and the temperature change in the sample is recorded. Usually, these measurements are performed to study protein stability with respect to denaturation. These studies give important insights into the conformational behavior of biological systems in the investigated temperature interval. Deviations from a constant heat absorption indicate changes in the intramolecular packing of the proteins or structural fluctuations and conformational rearrangements.[591, 609] The binding of a ligand to a protein influences its stability and heat capacity, which is expressed in pronounced changes of the DSC thermogram[610–613] and supports the interpretation of the interaction between the binding partners.[614–616]

5. Characterization and Evaluation of Ligand Binding with a Knowledge-Based Scoring Function

In the following, our own knowledge-based scoring function DrugScore will be briefly described. It is used to predict binding modes and affinities, and it can be applied to identify, in graphical terms, regions in the protein binding pocket that are favorable for interactions.[544, 546]

For the development of DrugScore, the structural information of 1376 crystallographically determined protein–ligand complexes was retrieved from the database ReliBase. Subsequently, this information was converted into statistical preferences based on 17 atom types.[545] The requirement to consider both specific interactions and entropy-dependent solvent contributions prompted us to use two terms: a distance-dependent atom–atom pair preference sampled up to 6 Å atom–atom distances (Figure 10), and a singlet preference dependent on the solvent-accessible surface of protein and ligand, both in the bound or unbound state. Definition of an appropriate reference state, to which the atom-type-specific distribution functions relate, is crucial for

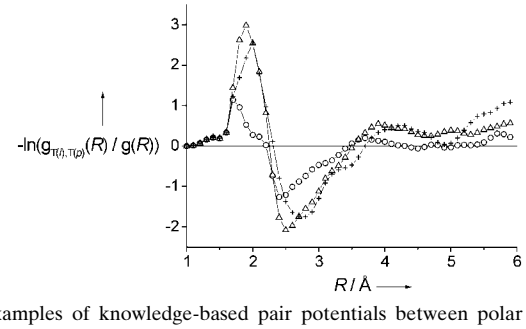


Figure 10. Examples of knowledge-based pair potentials between polar and charged (O.co2-N.pl3 (+), O.3-O.co2 (△), O.3-O.3 (○)) ligand and protein atoms as a function of distance. The first atom-type symbol refers to the ligand, the second to the protein.

the information stored in the respective preferences. In case of the pair preferences, a state based on a compact protein–ligand configuration with mean nonspecific interactions was selected. For the singlet preferences, a reference state with complete separation of protein and ligand was used. To evaluate a given protein–ligand binding mode, the various pair interactions and singlet contributions arising from all protein and ligand atoms were summed. Enthalpic and entropic contributions resulting from purely intramolecular effects were not considered.

Using an approximative grid-based method to compute the solvent-accessible surface, little computing time is required to evaluate each protein–ligand configuration. Because of the restriction to non-hydrogen atoms in the derivation of the pair preferences, no assumptions on possible protonation states are required using DrugScore. A triangular function is used for smoothing the pair and singlet distributions initially obtained from the structural data ("moving-window technique"). It should sufficiently "soften" the preferences to tolerate inherent deviations from ideal geometry caused by limited accuracy of protein crystal-structure analyses or limited precision of docked protein–ligand binding modes. As no protein- or ligand-type-specific training data set was used for the development of the function, its general applicability can be assumed.

To assess the reliability of DrugScore to identify near-native protein–ligand binding modes out of a number of clearly deviating geometries, data sets of 91 and 68 complexes were analyzed. Up to 500 protein–ligand configurations were produced with FlexX.[552] In 80% of the cases, the best-ranked ligand binding modes show an rms deviation of < 2.0 Å from the crystal-structure reference (native pose). With respect to the data set of 91 complexes, this corresponds to an improvement of 35% compared to the ranking obtained with the scoring function originally implemented in FlexX (Figure 11). Similarly, convincing results were obtained considering binding modes generated with DOCK.[566, 617]

In the context of a virtual screening study on human carbonic anhydrase II (hCAII),[618] the complex structures of two novel inhibitors (**10**, **11**; Scheme 4) discovered by means of a computer simulation were determined crystallographically. This allowed us to assess directly by experiment the binding modes which had been predicted in advance using DrugScore or the original scoring function in FlexX. As

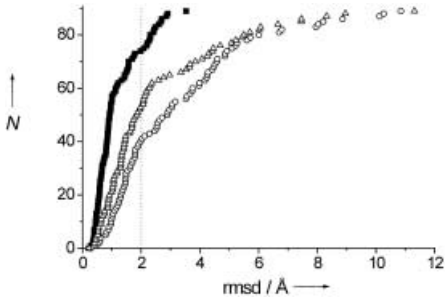
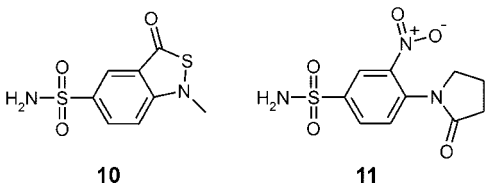


Figure 11. Accumulated number N of best-ranked docking solutions for 91 protein–ligand complexes as a function of the rms deviation with respect to the crystallographic reference structure. The ranking is based on the scoring function in FlexX (○) or DrugScore (△). For comparison, the accumulated number of complexes considering the best generated geometry, disregarding its actual rank, is plotted.(■). This distribution indicates the limit an ideal scoring function could achieve.



10 **11**

Scheme 4. Inhibitors discovered as a part of a virtual screening approach on hCAII.[618]

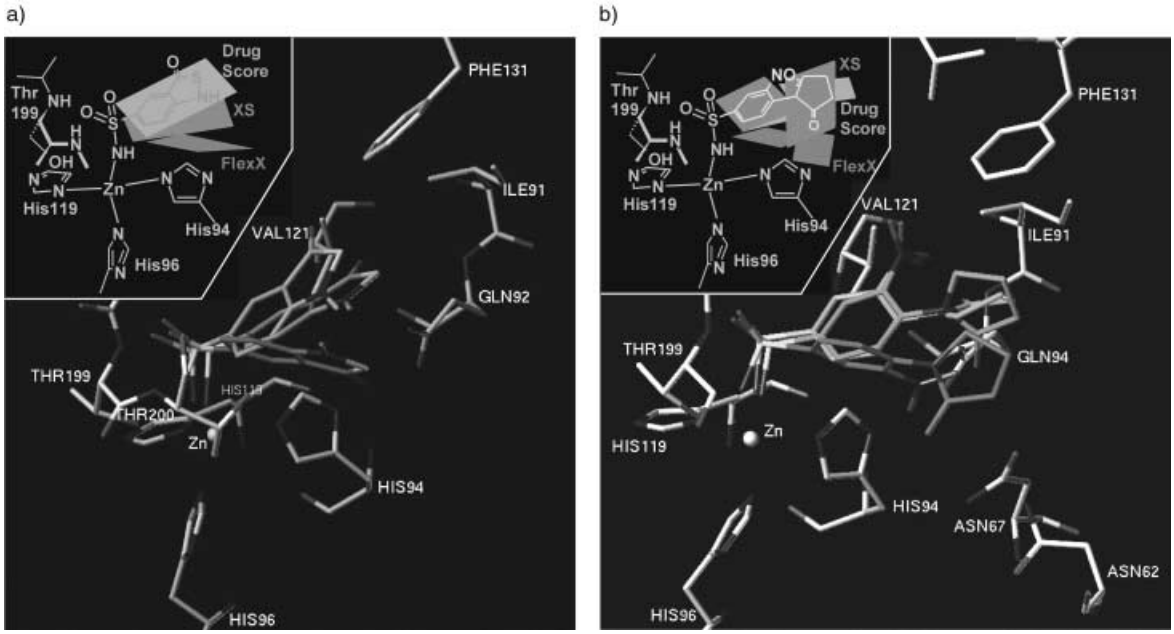


Figure 12. Superposition of the crystal structures (medium gray) and best-ranked docking solutions by DrugScore (light gray) and the scoring function in FlexX (dark gray) of two inhibitors [a): **10**, b): **11**]. They were discovered in virtual screening on hCAII.[618]

Figure 12 shows, the binding geometry predicted by DrugScore (rmsd value relative to the crystal structure: 1.2 Å for **10**; 1.4 Å for **11**) falls significantly closer to the experimentally observed geometry than does the solution proposed by FlexX, which suggests a configuration with an rmsd relative to the crystal structure of 2.2 Å for **10** and 2.7 Å for **11**. A reliable identification of near-native binding modes is not only of the utmost importance for subsequent ligand optimization, but also a prerequisite for the prediction of binding affinities. Importantly enough, such arrangements have to be produced by docking programs.

The validation of DrugScore with respect to affinity prediction was based on six data sets compiled from crystallographically determined protein–ligand complexes and three data sets with binding modes generated by FlexX. In the case of 16 serine protease–inhibitor complexes (X-ray structures), an r^2 value of 0.86 and a standard deviation of 0.95 logarithmic units was achieved compared to experimental affinities. For a set of 64 thrombin and trypsin inhibitors docked into the respective proteins, an r^2 value of 0.48 and a standard deviation of 0.71 log units was revealed. The deviations matched the experimental error.

"Hot spots" of binding can be calculated by using different probe atoms, using the distance-dependent pair preferences. The results can be visualized in terms of isocontour surfaces. They intuitively highlight favorable regions in the binding pocket suitable for a particular ligand-atom type (Figure 13 and frontispiece). They support ligand optimization and help to establish a protein-based pharmacophore hypothesis that can subsequently be used to screen for candidate molecules in compound libraries.

The prediction of hot spots has been validated using 159 protein–ligand complexes. Through the application of five probe atoms, the atom type observed in the crystal structure could be predicted correctly in 74% of the cases. Requesting only an atom type of appropriate physicochemical properties revealed correct predictions in 85% of the cases. Mapping the protein binding pocket of hCAII with GRID,[418] SuperStar,[556] LUDI,[619] and DrugScore, carried out as part of a virtual screening study,[618] demonstrated that all methods qualitatively identify the same regions in space, however with different relative weightings.

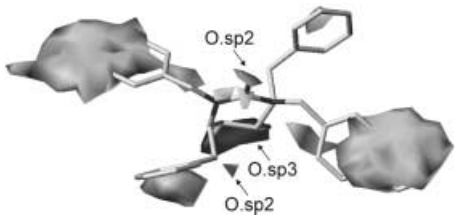
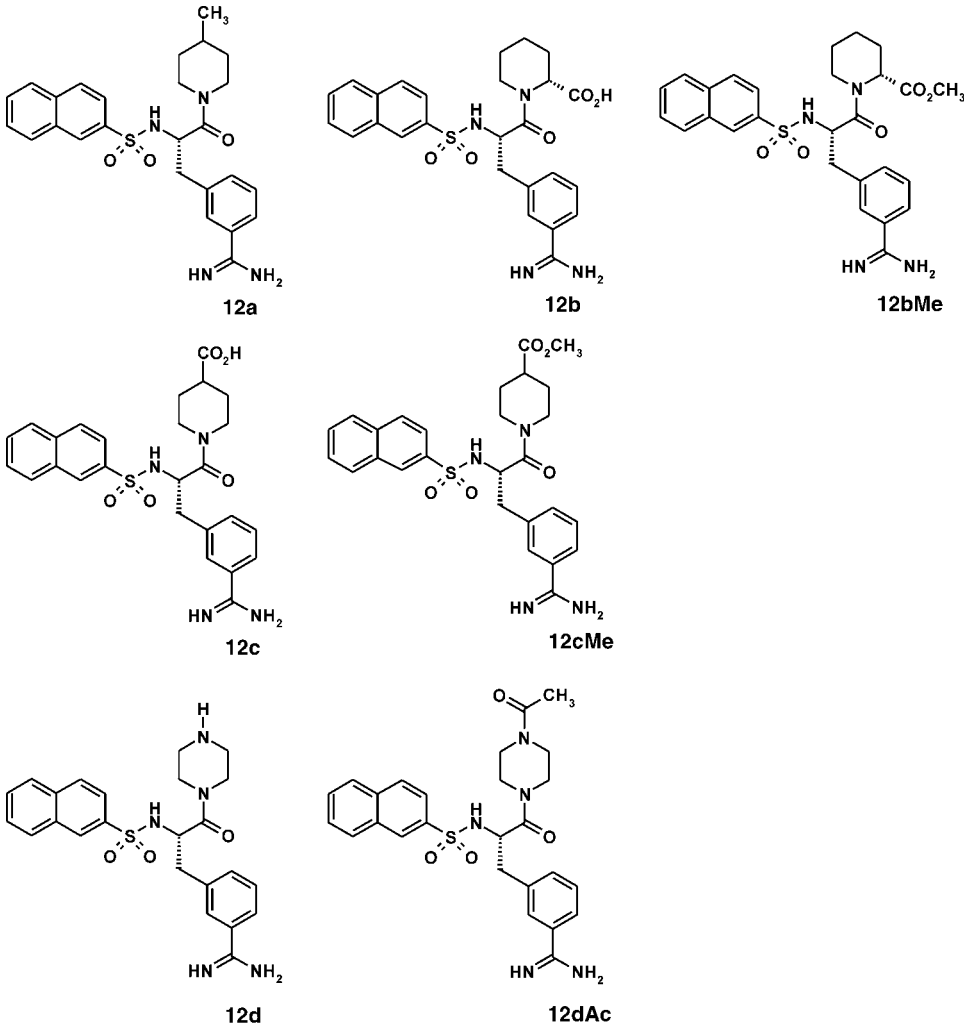


Figure 13. Hot spots in the binding pocket of HIV-1 protease (PDB code 1hvr) based on knowledge-based pair potentials in DrugScore. Values beyond a predefined threshold are isocontoured and shown together with the ligand XK263 from Merck for comparative purposes. Regions favorable for aromatic carbon atoms are colored in light gray, for carbonyl oxygen atoms in midgray, and for hydroxy groups in dark gray.

6. Factorization of Thermodynamic Contributions to Ligand Binding Using an Example of Serine Protease Inhibitors

Inhibitor binding to trypsin and thrombin, studied in our group, should serve as an example for the partitioning of affinity-determining factors.[607] A series of structurally related benzamidine inhibitors (**12a**–**12dAc**, Scheme 5) and the development compounds napsagatran (**13**), CRC220 (**14**),



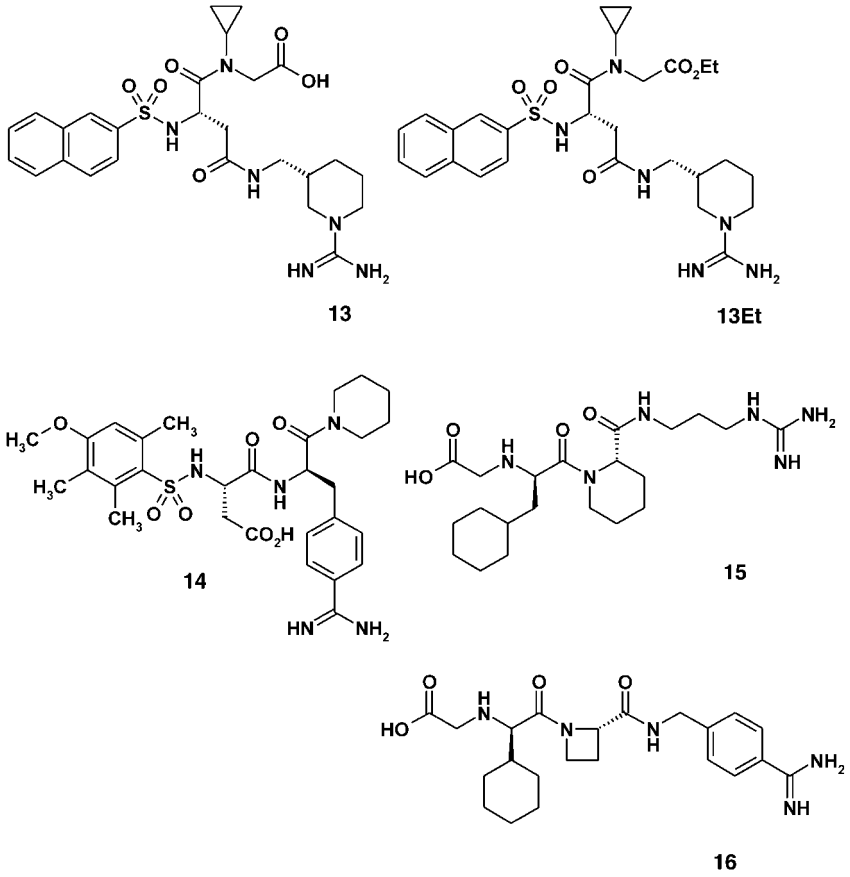
Scheme 5. Structurally related inhibitors of trypsin and thrombin.

inogatran (**15**), and melagatran (**16**; Scheme 6) have been investigated by crystallography and microcalorimetry.

Crystal-structure determination revealed uniform binding modes of all ligands with a salt bridge between the ligand amidino groups and Asp 189. Further interactions are formed between the central amide or sulfonamide groups and the protein carbonyl and amide NH groups of Gly 216 and 219. In detail, however, interesting deviations are observed that allow important conclusions with respect to thermodynamics.

Interestingly, the 2-carboxy derivative **12b** and napsagatran (**13**) acquires a proton upon protein binding, whereas the piperazine derivative **12d** releases a proton upon binding. The closely related 4-carboxy derivative **12c** remains deprotonated at the acid group during the binding process. The same observation holds for CRC220 (**14**); also here the acid group of the central aspartate remains deprotonated. The apparent pK_a shifts of the groups involved, which show very similar pK_a values in water, can be explained by examining binding modes of the different ligands. For CRC220, the carboxylate group is oriented away from the binding pocket towards the solvent (Figure 14); accordingly, this group remains partially solvated and the local dielectric conditions closely resemble those in the bulk water phase. For **12b** and napsagatran, the acid groups are protonated and point towards the catalytic serine. They obviously donate hydrogen bonds to the protein (Figure 14). The environment has such a strong influence on the local dielectric conditions (induced dielectric fit) that the pK_a values of the carboxy groups involved are shifted by more than four log units. The 4-carboxy group in **12b**, which does not change its deprotonated state upon binding, packs with parallel orientation next to the hydrophobic ring plane of the catalytic histidine. The induced pK_a shift caused by this environment is obviously not sufficient to protonate this acid group. The terminal amino group of **12d** with a pK_a value of 7.5 is partially protonated in the buffer medium of pH = 7.8, however, it binds in the deprotonated form. The hydrophobic protein environment reinforces a pK_a shift of this basic group towards smaller values.

The partitioning of enthalpic and entropic binding contributions for the free acids and esters **12b**/**12bMe** and **12c**/**12cMe** provides an instructive example. The inhibitors with the functional groups in the 4 position differ by more than 8 kJ mol^{-1} in ΔG° (**12c**: -35.5, **12cMe**: -43.6 kJ mol^{-1}), and the higher affinity of the ester is attributed to a stronger enthalpic binding (ΔH°; **12c**: -26.8, **12cMe**: -39.6 kJ mol^{-1}), whereas the entropic contributions $T\Delta S^\circ$ are comparable (**12c**: $+8.7$, **12cMe**: $+4.0$ kJ mol^{-1}). The enthalpically reduced binding of the acid at 25 °C can be explained by an unfavorable desolvation energy: the acid loses hydrogen-bonding partners for two polar acceptors (oxygen atoms), and is transferred from aqueous solution to the protein, whereas, the ester must only compensate for the desolvation of one carbonyl oxygen atom. The bridging ester oxygen atom exhibits practically no basic properties.[620] For the analogous pair with carboxylate and ester groups in the 2 position, a reverse correlation is observed. Here, acid and ester groups possess almost equal affinities (ΔG°; **12b**: -36.4, **12bMe**: -37.0 kJ mol^{-1}). However, the free acid is now enthalpically favored, the enthalpy contribution of the ester is significantly less exothermic (ΔH°; **12b**: -46.7, **12bMe**: -16.9 kJ mol^{-1}). For entropic reasons, the binding of the acid is now significantly less favorable ($T\Delta S^\circ$; **12b**: -10.6, **12bMe**: -20.6 kJ mol^{-1}). Upon binding, the protonated acid group in **12b** forms an enthalpically favored hydrogen bond to the protein (see above). At the same time, this part of the molecule loses residual mobility caused by immobilization by this additional hydrogen bond in the binding pocket. This fact results in a less favorable entropic contribution. The reduced temperature factors observed in the crystal structure for this part of the molecule **12b** further support this observation. The piperazine **12d** and the acetyl derivative **12dAc** exhibit very similar binding affinities (ΔG°; **12d**: -40.8, **12dAc**: -42.7 kJ mol^{-1}). After correcting for the superimposed deprotonation step of **12d**, similar contributions are observed for ΔH° and $T\Delta S^\circ$ in both cases (ΔH°; **12d**: -32.9, **12dAc**: -34.4, $T\Delta S^\circ$; **12d**: -7.9, **12dAc**: -8.2 kJ mol^{-1}). Both the free amine and the protected acetyl compound possess one polar atom in this group capable of forming hydrogen bonds. As in the bound state, no hydrogen bonds are formed to the protein in



Scheme 6. Development compounds which bind to trypsin and thrombin.

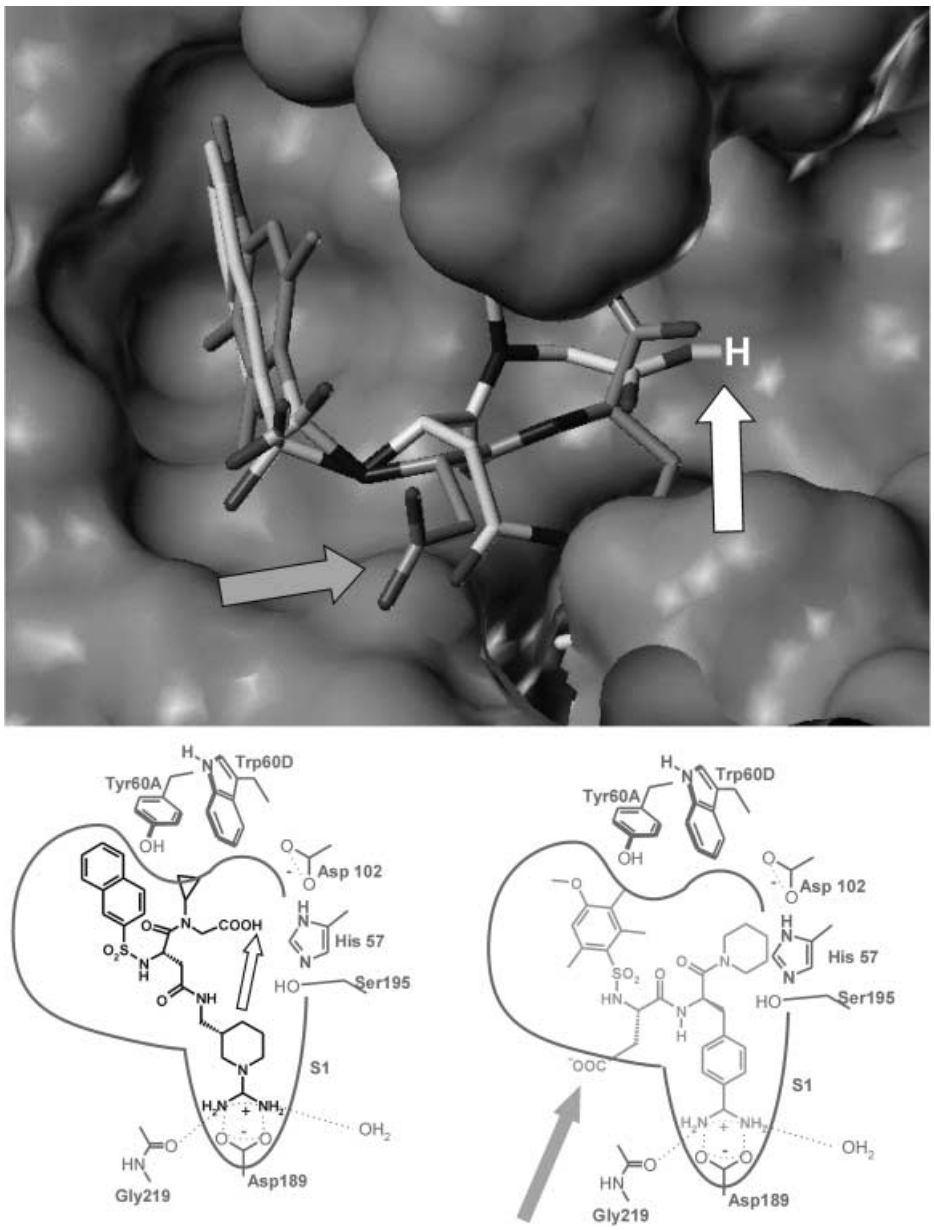


Figure 14. Schematic binding mode (bottom) of napsagatran (**13**) and CRC220 (**14**) to thrombin. The crystallographically determined binding geometries (light: napsagatran, white arrow indicates the protonated carboxylate group; dark: CRC220, gray arrow shows the deprotonated acid group) are superimposed.

either case; this results in an unfavorable desolvation, obviously with comparable thermodynamic contributions for both groups. Additional compounds investigated in this series confirm the discussed trends.

A comparison of the heat-capacity changes is equally interesting. In all cases, a strongly negative ΔC_p is observed, which means that ΔH° becomes increasingly exothermic with rising temperature and, as ΔG° is essentially temperature independent, the binding process becomes entropically less favorable. The surface portions that are buried upon binding were derived from the known crystal structures of the corresponding ligand–protein complexes. As thrombin and trypsin are described as relatively rigid proteins, the surface portions of the uncomplexed protein were calculated using the structure of the protein simply by removing the bound ligand. If the empirical correlation derived from protein folding (see above) is applied to estimate ΔC_p, the ΔC_p values predicted are too low. Obviously, smaller heat capacities are observed experimentally for the complexes than expected from solvation-dependent surface contributions alone. Presumably, the additional factors described above must be considered for a detailed structural interpretation of ΔC_p. Surprisingly, in contrast to trypsin, a sodium-specific dependence of ΔC_p is found for thrombin. An according allosteric regulation of thrombin is known, however, this ion-specific effect cannot yet be explained in structural terms.

7. Summary and Outlook

The affinity of a small-molecule ligand for a macromolecular receptor usually serves as a criterion to define the biological activity of the respective compound. It is determined by electrostatic interactions between the ligand and the receptor, together with contributions from solvation and desolvation, and the spatial complementarity of both binding partners. Additional influences arise from changes in the number of degrees of freedom and conformational changes of ligand and receptor experienced upon complex formation. Understanding of the determining enthalpic and entropic contributions to binding is a prerequisite for affinity predictions.

The methods of virtual screening and rational drug design require rapid and reliable methods for affinity prediction. In this contribution we have described and classified known theoretical approaches with respect to their methodological foundations. Significant differences in computational requirements and the general scope of the methods have been addressed. The pragmatic combination of several scoring methods as in the so-called consensus approaches indicates that, at present, no general-purpose method is available that adequately considers all above-described relevant contributions to ligand-receptor binding. However, advantages have recently been achieved for the (rapid) prediction of binding affinity by means of newly developed knowledge-based scoring functions.

Physicochemical techniques to quantitatively characterize ligand–receptor binding have been developed and further

enhanced in recent years. In particular, direct access to thermodynamic parameters by means of microcalorimetry provides new insights into the thermodynamic foundations of binding affinity. Supported by the structural characterization of ligand–receptor complexes, it is possible to factorize binding affinity into individual contributions. However, the described examples of serine-protease inhibitors demonstrate that steps superimposed on the binding process must be considered and adequately handled.

With the increase in structural information and binding data about receptor–ligand complexes, further advances in the understanding and the description of binding affinity can be expected. This will improve our methods for its prediction. In particular, current approximations—such as the complete disregard of changes in protonation state of the binding partners upon complex formation, the consideration of receptors as rigid entities, the neglect of allosteric effects, and the to date inadequate handling of water molecules in ligand binding—will be the starting points for new developments.

Abbreviation list

ASP	atomic solvation parameters
CSD	Cambridge Structural Databank, a database of crystal structures of small molecules[1]
FEP–MD	free-energy-perturbation calculations/molecular dynamics
GB/SA	generalized Born approach
LIE	linear interaction energy
ME	master equation
NBTI	non-Boltzmann thermodynamic integration
PBE	Poisson–Boltzmann equation
PDB	protein database[2]
PLS	partial least-squares method
QSAR	quantitative structure–activity relationships
rmsd	root-mean-square deviation in the Cartesian coordinates of mutually corresponding atoms in two molecules

Received: 22nd October 2001 [A 497]

[1] F. H. Allen, J. E. Davies, J. J. Galloy, O. Johnson, O. Kennard, C. F. Macrae, E. M. Mitchell, G. F. Mitchell, J. M. Smith, D. G. Watson, *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187–204.

[2] F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. E. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, *J. Mol. Biol.* **1977**, *112*, 535–542.

[3] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993.

[4] P. Ehrlich, *Lancet* **1913**, *II*, 445–451.

[5] D. E. Koshland, *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 98–104.

[6] D. E. Koshland, *Angew. Chem.* **1994**, *106*, 2468–2472; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2375–2378.

[7] J. P. Glusker, M. Lewis, M. Rossi, *Crystal Structure Analysis for Chemists and Biologists*, VCH, Weinheim, **1994**.

[8] J. Drenth, *Principles of Protein X-ray Crystallography*, Springer, New York, **1999**.

[9] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, **1986**.

[10] G. Siegal, J. van Duynhoven, M. Baldus, *Curr. Opin. Chem. Biol.* **1999**, *3*, 530–536.

[11] G. M. Clore, A. M. Gronenborn, *Science* **1991**, *252*, 1390–1399.

[12] W. Kühlbrandt, K. A. Williams, *Curr. Opin. Chem. Biol.* **1999**, *3*, 537–543.

[13] M. H. Stowell, A. Miyazawa, N. Unwin, *Curr. Opin. Struct. Biol.* **1998**, *8*, 595–600.

[14] W. Chiu, A. McGough, M. B. Sherman, M. F. Schmid, *Trends Cell Biol.* **1999**, *9*, 154–159.

[15] W. Baumeister, C. S. Alasdair, *Trends Biochem. Sci.* **2000**, *25*, 624–631.

[16] F. S. Collins, A. Patrinos, E. Jordan, A. Chakravarti, R. Gesteland, L. Walters, E. Fearon, L. Hartwelt, C. H. Langley, R. A. Mathies, M. Olson, A. J. Pawson, T. Pollard, A. Williamson, B. Wold, K. Buetow, E. Branscomb, M. Capecchi, G. Church, H. Garner, R. A. Gibbs, T. Hawkins, K. Hodgson, M. Knotek, M. Meisler, G. M. Rubin, L. M. Smith, M. Westerfield, E. W. Clayton, N. L. Fisher, C. E. Lerman, J. D. McInerney, W. Nebo, N. Press, D. Valle, *Science* **1998**, *282*, 682–689.

[17] G. M. Rubin, M. D. Yandell, J. R. Wortman, G. L. Gabor Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, P. W. Li, R. Apweiler, W. Fleischmann, J. M. Cherry, S. Henikoff, M. P. Skupski, S. Misra, M. Ashburner, E. Birney, M. S. Boguski, T. Brody, P. Brokstein, S. E. Celniker, S. A. Chervitz, D. Coates, A. Cravchik, A. Gabrielian, R. F. Galle, W. M. Gelbart, R. A. George, L. S. Goldstein, F. Gong, P. Guan, N. L. Harris, B. A. Hay, R. A. Hoskins, J. Li, Z. Li, R. O. Hynes, S. J. Jones, P. M. Kuehl, B. Lemaitre, J. T. Littleton, D. K. Morrison, C. Mungall, P. H. O'Farrell, O. K. Pickeral, C. Shue, L. B. Vosshall, J. Zhang, Q. Zhao, X. H. Zheng, F. Zhong, W. Zhong, R. Gibbs, J. C. Venter, M. D. Adams, S. Lewis, *Science* **2000**, *287*, 2204–2215.

[18] S. Broder, J. C. Venter, *Annu. Rev. Pharmacol. Toxicol.* **2000**, *40*, 97–132.

[19] See edition "The Human Genome": *Nature* **2001**, *409*, 813–958.

[20] E. M. Marcotte, M. Pellegrini, H. Ng, D. W. Rice, T. O. Yeates, D. Eisenberg, *Science* **1999**, *285*, 751–753.

[21] F. Lottspeich, *Angew. Chem.* **1999**, *111*, 2630–2647; *Angew. Chem. Int. Ed.* **1999**, *38*, 2477–2492.

[22] J. H. Wang, R. M. Hewick, *Drug Discovery Today* **1999**, *4*, 129–383.

[23] A. Danchin, *Curr. Opin. Struct. Biol.* **1999**, *9*, 363–367.

[24] C. A. Orengo, A. E. Todd, J. M. Thornton, *Curr. Opin. Struct. Biol.* **1999**, *9*, 374–382.

[25] D. R. Westhead, J. M. Thornton, *Curr. Opin. Biotechnol.* **1998**, *9*, 383–389.

[26] B. Rost, *Structure* **1998**, *6*, 259–263.

[27] S. K. Burley, S. C. Almo, J. B. Bonanno, M. Capel, M. R. Chance, T. Gaasterland, D. Lin, A. Sali, F. W. Studier, S. Swaminathan, *Nat. Genet.* **1999**, *23*, 151–157.

[28] I. Uson, G. M. Sheldrick, *Curr. Opin. Struct. Biol.* **1999**, *9*, 643–648.

[29] D. J. Diller, M. R. Redinbo, E. Pohl, W. G. J. Hol, *Proteins* **1999**, *36*, 526–541.

[30] D. J. Diller, E. Pohl, M. R. Redinbo, B. T. Hovey, W. G. J. Hol, *Proteins* **1999**, *36*, 512–525.

[31] D. A. Jones, F. A. Fitzpatrick, *Curr. Opin. Chem. Biol.* **1999**, *3*, 71–76.

[32] J. Skolnick, J. S. Fetrow, *Trends Biotechnol.* **2000**, *18*, 34–39.

[33] F. Spaltmann, M. Blunck, K. Ziegelbauer, *Drug. Discovery Today* **1999**, *4*, 17–26.

[34] J. Drews in *Human Disease—from Genetic Causes to Biochemical Effects* (Eds.: J. Drews, S. Ryser), Blackwell, Berlin, **1997**, pp. 5–9.

[35] G. A. Petsco, *Nature* **1996**, *384*, 7–9.

[36] J. Drews, *Drug Discovery Today* **2000**, *5*, 2–4.

[37] J. Drews, *Nat. Biotechnol.* **1996**, *14*, 1516–1518.

[38] R. B. Silvermann, *Medizinische Chemie für Organiker*, *Biochemiker und Pharmazeutische Chemiker*, VCH, Weinheim, **1994**.

[39] H.-J. Böhm, G. Klebe, H. Kubinyi, *Wirkstoffdesign*, Spektrum, Heidelberg, **1996**.

[40] O. H. Chan, B. H. Stewart, *Drug Discovery Today* **1996**, *1*, 461–473.

[41] D. A. Smith, H. van de Waterbeemd, *Curr. Opin. Chem. Biol.* **1999**, *3*, 373–378.

[42] A. P. Watt, D. Morrison, D. C. Evans, *Drug Discovery Today* **2000**, *5*, 17–24.

[43] D. B. Searls, *Drug Discovery Today* **2000**, *5*, 135 – 143.
[44] G. R. Lenz, H. M. Nash, S. Jindal, *Drug. Discovery Today* **2000**, *5*, 145 – 156.
[45] M. H. Beydon, A. Fournier, L. Drugeault, J. Becquart, *J. Biomol. Screening* **2000**, *5*, 13 – 22.
[46] R. P. Hertzberg, A. J. Pope, *Curr. Opin. Chem. Biol.* **2000**, *4*, 445 – 451.
[47] W. P. Walters, M. T. Stahl, M. A. Murcko, *Drug Discovery Today* **1998**, *3*, 160 – 178.
[48] *Virtual Screening for Bioactive Molecules*, *Vol. 10* (Eds.: H.-J. Böhm, G. Schneider), Wiley-VCH, Weinheim, **2000**.
[49] I. D. Kuntz, *Science* **1992**, *257*, 1078 – 1082.
[50] I. D. Kuntz, E. C. Meng, B. K. Shoichet, *Acc. Chem. Res.* **1994**, *27*, 117 – 123.
[51] D. B. Boyd, *Mod. Drug Discovery* **1998**, *6*, 41 – 48.
[52] E. M. Gordon, M. A. Gallop, D. V. Patel, *Acc. Chem. Res.* **1996**, *29*, 144 – 154.
[53] T. Carell, E. A. Wintner, J. Rebek, Jr., *Angew. Chem.* **1994**, *106*, 2162 – 2164; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2061 – 2064.
[54] J. G. Houston, M. Banks, *Curr. Opin. Biotechnol.* **1997**, *8*, 734 – 740.
[55] N. K. Terrett, M. Gardner, D. W. Gordon, R. J. Kobylecki, J. Steele, *Tetrahedron* **1995**, *51*, 8135 – 8173.
[56] F. Balkenhohl, C. von dem Bussche-Hünnefeld, A. Lansky, C. Zechel, *Angew. Chem.* **1996**, *108*, 2436 – 2488; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2289 – 2337.
[57] S. Kauffman, A. D. Ellington, *Curr. Opin. Chem. Biol.* **1999**, *3*, 256 – 259.
[58] A. Dömling, I. Ugi, *Angew. Chem.* **2000**, *112*, 3300 – 3344; *Angew. Chem. Int. Ed.* **2000**, *39*, 3169 – 3210.
[59] S. F. Oliver, C. Abell, *Curr. Opin. Chem. Biol.* **1999**, *3*, 299 – 306.
[60] R. Lahana, *Drug Discovery Today* **1999**, *4*, 447 – 448.
[61] M. A. Sills, *Drug Discovery Today* **1998**, *3*, 304 – 312.
[62] W. A. Warr, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 134 – 140.
[63] D. E. Clark, S. D. Pickett, *Drug Discovery Today* **2000**, *5*, 49 – 58.
[64] J. F. Blake, *Curr. Opin. Biotechnol.* **2000**, *11*, 104 – 107.
[65] J. H. Van Drie, M. S. Lajiness, *Drug Discovery Today* **1998**, *3*, 274 – 283.
[66] L. Weber, S. Wallbaum, C. Broger, K. Gubernator, *Angew. Chem.* **1995**, *107*, 2452 – 2454; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2280 – 2282.
[67] C. M. Murray, S. J. Cato, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 46 – 50.
[68] S. J. Teague, A. M. Davis, P. D. Leeson, T. Oprea, *Angew. Chem.* **1999**, *111*, 3962 – 3967; *Angew. Chem. Int. Ed.* **1999**, *38*, 3743 – 3748.
[69] S. L. Schreiber, *Science* **2000**, *287*, 1964 – 1969.
[70] *Designing Bioactive Molecules* (Eds.: Y. C. Martin, P. Willet, S. R. Heller), American Chemical Society, Washington, **1999**.
[71] H.-J. Böhm, M. Böhringer, D. Bur, H. Gmuender, W. Huber, W. Klaus, D. Kostrewa, H. Kühne, T. Lübbers, N. Meunier-Keller, F. Müller, *J. Med. Chem.* **2000**, *43*, 2664 – 2674.
[72] J. M. Chen, F. C. Nelson, J. I. Levin, D. Mobilio, F. J. Moy, R. Nilakantan, A. Zask, R. Powers, *J. Am. Chem. Soc.* **2000**, *122*, 9648 – 9654.
[73] B. Buzbee, *Science* **1993**, *261*, 852 – 853.
[74] J. Couzin, *Science* **1998**, *281*, 762.
[75] G. Klebe, *Period. Biol.* **1998**, *100 (Suppl. 2)*, 93 – 98.
[76] H. Kubinyi, *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 4 – 15.
[77] J. Greer, J. W. Erickson, J. J. Baldwin, M. D. Varney, *J. Med. Chem.* **1994**, *37*, 1035 – 1054.
[78] D. B. Boyd, *Rev. Comput. Chem.* **1990**, *1*, 355 – 371.
[79] R. E. Babine, S. L. Bender, *Chem. Rev.* **1997**, *97*, 1359 – 1472.
[80] D. H. Harrison, K. M. Bohren, G. A. Petsko, D. Ringe, K. H. Gabbay, *Biochemistry* **1997**, *36*, 16134 – 16140.
[81] L. Costantino, G. Rastelli, K. Vescovini, G. Cignarella, P. Vianello, A. D. Corso, M. Cappiello, U. Mura, D. Barlocco, *J. Med. Chem.* **1996**, *39*, 4396 – 4405.
[82] I. R. Hardcastle, M. G. Rowlands, J. Houghton, I. B. Parr, G. A. Potter, M. Jarmann, K. J. Edwards, C. A. Laughton, J. O. Trent, S. Neidle, *J. Med. Chem.* **1995**, *38*, 241 – 248.
[83] A. Scozzafava, C. T. Supuran, *J. Med. Chem.* **2000**, *43*, 3677 – 3687.
[84] P. A. Boriack-Sjodin, S. Zeitlin, H. H. Chen, L. Crenshaw, S. Gross, A. Dantanarayana, P. Delgado, J. A. May, T. Dean, D. W. Christianson, *Protein Sci.* **1998**, *7*, 2483 – 2489.
[85] L. R. Scolnick, A. M. Clements, J. Liao, L. Crenshaw, M. Hellberg, J. May, T. R. Dean, D. W. Christianson, *J. Am. Chem. Soc.* **1997**, *119*, 850 – 851.
[86] A. Jain, G. M. Whitesides, R. S. Alexander, D. W. Christianson, *J. Med. Chem.* **1994**, *37*, 2100 – 2105.
[87] A. S. Kalgutkar, B. C. Crews, S. W. Rowlinson, A. B. Marnett, K. R. Kozak, R. P. Remmel, L. J. Marnett, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 925 – 930.
[88] C. I. Bayly, W. C. Black, S. Leger, N. Ouimet, M. Ouellet, M. D. Percival, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 307 – 312.
[89] C. A. Veale, J. R. Damewood, G. B. Steelman, C. Bryant, B. Gomes, J. Williams, *J. Med. Chem.* **1995**, *38*, 86 – 97.
[90] R. J. Cregge, S. L. Durham, R. A. Farr, S. L. Gallion, C. M. Hare, R. V. Hoffman, M. J. Janusz, H. O. Kim, J. R. Koehl, S. Mehdi, W. A. Metz, N. P. Peet, J. T. Pelton, H. A. Schreuder, S. Sunder, C. Tardif, *J. Med. Chem.* **1998**, *41*, 2461 – 2480.
[91] P. R. Bernstein, D. Andisik, P. K. Bradley, C. B. Bryant, C. Ceccarelli, J. R. Damewood, R. Earley, P. D. Edwards, S. Feeney, B. C. Gomes, B. J. Kosmider, G. B. Steelman, R. M. Thomas, E. P. Vacek, C. A. Veale, J. C. Williams, D. J. Wolanin, S. A. Woolson, *J. Med. Chem.* **1994**, *37*, 3313 – 3326.
[92] M. B. Andrus, S. L. Schreiber, *J. Am. Chem. Soc.* **1993**, *115*, 10420 – 10421.
[93] P. S. Dragovich, J. E. Barker, J. French, M. Imbacuan, V. J. Kalish, C. R. Kissinger, D. R. Knighton, C. T. Lewis, E. W. Moomaw, H. E. Parge, L. A. Pelletier, T. J. Prins, R. E. Showalter, J. H. Tatlock, K. D. Tucker, J. E. Villafranca, *J. Med. Chem.* **1996**, *39*, 1872 – 1884.
[94] T. Lubbers, P. Angehrn, H. Gmunder, S. Herzig, J. Kulhanek, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 821 – 826.
[95] A. Tossi, I. Bonin, N. Antcheva, S. Norbedo, F. Benedetti, S. Miertus, A. C. Nair, T. Maliar, F. Dal Bello, G. Palu, D. Romeo, *Eur. J. Biochem.* **2000**, *267*, 1715 – 1722.
[96] A. K. Ghosh, K. Krishnan, D. E. Walters, W. Cho, H. Cho, Y. Koo, J. Trevino, L. Holland, J. Buthod, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 979 – 982.
[97] T. Y. Lee, V. D. Le, D. Y. Lim, Y. C. Lin, G. M. Morris, A. L. Wong, A. J. Olson, J. H. Elder, C. H. Wong, *J. Am. Chem. Soc.* **1999**, *121*, 1145 – 1155.
[98] H. I. Skulnick, P. D. Johnson, P. A. Aristoff, J. K. Morris, K. D. Lovasz, W. J. Howe, K. D. Watenpaugh, M. N. Janakiraman, D. J. Anderson, R. J. Reischer, T. M. Schwartz, L. S. Banitt, P. K. Tomich, J. C. Lynn, M.-M. Horng, K.-T. Chong, R. R. Hinshaw, L. A. Dolak, E. P. Seest, F. J. Schwende, B. D. Rush, G. M. Howard, L. N. Toth, K. R. Wilkinson, T. J. Kakuk, C. W. Johnson, S. L. Cole, R. M. Zaya, G. L. Zipp, P. L. Possert, R. J. Dalga, W.-Z. Zhong, M. G. Williams, K. R. Romines, *J. Med. Chem.* **1997**, *40*, 1149 – 1164.
[99] G. V. DeLucca, S. Erickson Viitanen, P. Y. S. Lam, *Drug Discovery Today* **1997**, *2*, 6 – 18.
[100] S. Thaisrivongs, D. L. Romero, R. A. Tommasi, M. N. Janakiraman, J. W. Strohbach, S. R. Turner, C. Biles, R. R. Morge, P. D. Johnson, P. A. Aristoff, P. K. Tomich, J. C. Lynn, M. M. Horng, K. T. Chong, R. R. Hinshaw, W. J. Howe, B. C. Finzel, K. D. Watenpaugh, *J. Med. Chem.* **1996**, *39*, 4630 – 4642.
[101] S. H. Reich, M. Melnick, M. J. Pino, M. A. Fuhry, A. J. Trippe, K. Appelt, J. F. Davies II, B. W. Wu, L. Musick, *J. Med. Chem.* **1996**, *39*, 2781 – 2794.
[102] M. D. Varney, K. Appelt, V. Kalish, M. R. Reddy, J. Tatlock, C. L. Palmer, W. H. Romines, B. W. Wu, L. Musick, *J. Med. Chem.* **1994**, *37*, 2274 – 2284.
[103] M. A. Navia, R. D. Tung, P. R. Chaturvedi, B. G. Rao, J. A. Partaledis, E. E. Kim, *FASEB J.* **1996**, *10*, 2436.
[104] H. M. Cheng, P. Keitz, J. B. Jones, *J. Org. Chem.* **1994**, *59*, 7671 – 7676.
[105] J. A. Montgomery, S. Niwas, J. D. Rose, J. A. Secrist III, Y. S. Babu, C. E. Bugg, M. D. Erion, W. C. Guida, S. E. Ealick, *J. Med. Chem.* **1993**, *36*, 55 – 69.
[106] M. S. Plummer, A. Shahripour, J. S. Kaltenbronn, E. A. Lunney, B. A. Steinbaugh, J. M. Hamby, H. W. Hamilton, T. K. Sawyer, C. Humblet, A. M. Doherty, M. D. Taylor, G. Hingorani, B. L. Batley, S. T. Rapundalo, *J. Med. Chem.* **1995**, *38*, 2893 – 2905.
[107] J. Rahuel, V. Rasetti, J. Maibaum, H. Rueger, R. Goschke, N. C. Cohen, S. Stutz, F. Cumin, W. Fuhrer, J. M. Wood, M. G. Grutter, *Chem. Biol.* **2000**, *7*, 493 – 504.

[108] B. A. Lefker, W. A. Hada, A. S. Wright, W. H. Martin, I. A. Stock, G. K. Schulte, J. Pandit, D. E. Danley, M. J. Ammirati, S. F. Sneddon, *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2623 – 2626.
[109] A. B. Smith III, R. Akaishi, D. R. Jones, T. P. Keenan, M. C. Guzman, R. C. Holcomb, P. A. Sprengeler, J. L. Wood, R. Hirschmann, M. K. Holloway, *Biopolymers* **1995**, *37*, 29 – 53.
[110] C. Mao, E. A. Sudbeck, T. K. Venkatachalam, F. M. Uckun, *Biochem. Pharmacol.* **2000**, *60*, 1251 – 1265.
[111] R. Silvestri, M. Artico, G. De Martino, E. Novellino, G. Greco, A. Lavecchia, S. Massa, A. G. Loi, S. Doratiotto, P. La Colla, *Bioorg. Med. Chem.* **2000**, *8*, 2305 – 2309.
[112] M. Artico, R. Silvestri, E. Pagnozzi, B. Bruno, E. Novellino, G. Greco, S. Massa, A. Ettorre, A. G. Loi, F. Scintu, P. La Colla, *J. Med. Chem.* **2000**, *43*, 1886 – 1891.
[113] C. Mao, E. A. Sudbeck, T. K. Venkatachalam, F. M. Uckun, *Antiviral Chem. Chemother.* **1999**, *10*, 233 – 240.
[114] E. A. Sudbeck, C. Mao, R. Vig, T. K. Venkatachalam, L. Tuel-Ahlgren, F. M. Uckun, *Antimicrob. Agents Chemother.* **1998**, *42*, 3225 – 3233.
[115] R. Vig, C. Mao, T. K. Venkatachalam, L. Tuel-Ahlgren, E. A. Sudbeck, F. M. Uckun, *Bioorg. Med. Chem.* **1998**, *6*, 1789 – 1797.
[116] I. Kang, J. H. Wang, *J. Biol. Chem.* **1994**, *269*, 12 024 – 12 031.
[117] D. W. Zaharevitz, R. Gussio, A. Wiegand, R. Jalluri, N. Pattabiraman, G. E. Kellogg, L. A. Pallansch, S. S. Yang, R. W. Buckheit, *Med. Chem. Res.* **1999**, *9*, 551 – 564.
[118] T. P. Kogan, B. Dupré, K. M. Keller, I. L. Scott, H. Bui, R. V. Market, P. J. Beck, J. A. Voytus, B. M. Revelle, D. Scott, *J. Med. Chem.* **1995**, *38*, 4976 – 4984.
[119] C. M. Huwe, T. J. Woltering, J. Jiricek, G. Weitz-Schmidt, C. H. Wong, *Bioorg. Med. Chem.* **1999**, *7*, 773 – 788.
[120] C. H. Wong, F. Moris-Varas, S.-C. Hung, T. G. Marron, C.-C. Lin, K. W. Gong, G. Weitz-Schmidt, *J. Am. Chem. Soc.* **1997**, *119*, 8152 – 8158.
[121] M. von Itzstein, P. Colman, *Curr. Opin. Struct. Biol.* **1996**, *6*, 703 – 709.
[122] M. von Itzstein, J. C. Dyason, S. W. Oliver, H. F. White, W.-Y. Wu, G. B. Kok, M. S. Pegg, *J. Med. Chem.* **1996**, *39*, 388 – 391.
[123] J. C. Wilson, R. J. Thomson, J. C. Dyason, P. Florio, K. J. Quelch, S. Abo, M. von Itzstein, *Tetrahedron: Asymmetry* **2000**, *11*, 53 – 73.
[124] N. R. Taylor, A. Cleasby, O. Singh, T. Skarzynski, A. J. Wonacott, P. W. Smith, S. L. Sollis, P. D. Howes, P. C. Cherry, R. Bethell, P. Colman, J. Varghese, *J. Med. Chem.* **1998**, *41*, 798 – 807.
[125] P. W. Smith, S. L. Sollis, P. D. Howes, P. C. Cherry, I. D. Starkey, K. N. Cobley, H. Weston, J. Scicinski, A. Merritt, A. Whittington, P. Wyatt, N. Taylor, D. Green, R. Bethell, S. Madar, R. J. Fenton, P. J. Morley, T. Pateman, A. Beresford, *J. Med. Chem.* **1998**, *41*, 787 – 797.
[126] G. Taylor, *Curr. Opin. Struct. Biol.* **1996**, *6*, 830 – 837.
[127] S. Singh, M. J. Jedrzejas, G. M. Air, M. Luo, W. G. Laver, W. J. Brouillette, *J. Med. Chem.* **1995**, *38*, 3217 – 3225.
[128] W. Lew, X. Chen, C. U. Kim, *Curr. Med. Chem.* **2000**, *7*, 663 – 672.
[129] V. R. Atigadda, W. J. Brouillette, F. Duarte, Y. S. Babu, S. Bantia, P. Chand, N. Chu, J. A. Montgomery, D. A. Walsh, E. Sudbeck, J. Finley, G. M. Air, M. Luo, G. W. Laver, *Bioorg. Med. Chem.* **1999**, *7*, 2487 – 2497.
[130] J. B. Finley, V. R. Atigadda, F. Duarte, J. J. Zhao, W. J. Brouillette, G. M. Air, M. Luo, *J. Mol. Biol.* **1999**, *293*, 1107 – 1119.
[131] J. N. Varghese, P. W. Smith, S. L. Sollis, T. J. Blick, A. Sahasrabudhe, J. L. McKimm-Breschkin, P. M. Colman, *Structure* **1998**, *6*, 735 – 746.
[132] C. U. Kim, W. Lew, M. A. Williams, H. T. Liu, L. J. Zhang, S. Swaminathan, N. Bischofberger, M. S. Chen, D. B. Mendel, C. Y. Tai, W. G. Laver, R. C. Stevens, *J. Am. Chem. Soc.* **1997**, *119*, 681 – 690.
[133] B. A. Katz, C. R. Johnson, R. T. Cass, *J. Am. Chem. Soc.* **1995**, *117*, 8541 – 8547.
[134] P. C. Weber, M. W. Pantoliano, D. M. Simons, F. R. Salemme, *J. Am. Chem. Soc.* **1994**, *116*, 2717 – 2727.
[135] B. A. Katz, *Biomol. Eng.* **1999**, *16*, 57 – 65.
[136] B. A. Katz, B. Liu, R. Cass, *J. Am. Chem. Soc.* **1996**, *118*, 7914 – 7920.
[137] B. P. Morgan, D. R. Holland, B. W. Matthews, P. A. Bartlett, *J. Am. Chem. Soc.* **1994**, *116*, 3251 – 3260.
[138] K. Hilpert, J. Ackermann, D. W. Banner, A. Gast, K. Gubernator, P. Hadvary, L. Labler, K. Müller, G. Schmid, T. B. Tschopp, H. van de Waterbeemd, *J. Med. Chem.* **1994**, *37*, 3889 – 3901.
[139] J. J. Slon-Usakiewicz, J. Sivaraman, Y. Li, M. Cygler, Y. Konishi, *Biochemistry* **2000**, *39*, 2384 – 2391.
[140] S. Hanessian, E. Balaux, D. Musil, L. L. Olsson, I. Nilsson, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 243 – 247.
[141] M. Li, Z. Lin, M. E. Johnson, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1957 – 1962.
[142] H. J. Böhm, D. W. Banner, L. Weber, *J. Comput. Aided Mol. Des.* **1999**, *13*, 51 – 56.
[143] J. Wagner, J. Kallen, C. Ehrhardt, J. P. Evenou, D. Wagner, *J. Med. Chem.* **1998**, *41*, 3664 – 3674.
[144] P. Burkhard, P. Taylor, M. D. Walkinshaw, *J. Mol. Biol.* **1998**, *277*, 449 – 466.
[145] U. Obst, D. W. Banner, L. Weber, F. Diederich, *Chem. Biol.* **1997**, *4*, 287 – 295.
[146] B. Shoichet, R. Stroud, D. Santi, I. Kuntz, K. Perry, *Science* **1993**, *259*, 1445 – 1450.
[147] *The Practice of Medicinal Chemistry* (Ed.: C. G. Wermuth), Academic Press, London, **1996**.
[148] J. P. Vacca, J. H. Condra, *Drug Discovery Today* **1997**, *2*, 261 – 272.
[149] P. Wutzler, *Infection* **2000**, *28*, 261 – 266.
[150] *3D QSAR in Drug Design: Theory, Methods and Applications* (Ed.: H. Kubinyi), Escom, Leiden, **1993**.
[151] *3D QSAR in Drug Design: Recent Advances* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/Escom, Leiden, **1997**.
[152] T. I. Oprea, C. L. Waller, *Rev. Comput. Chem.* **1997**, *11*, 127 – 182.
[153] G. Greco, E. Novellino, Y. C. Martin, *Rev. Comput. Chem.* **1997**, *11*, 183 – 240.
[154] *Chemometric Methods in Molecular Design*, *Vol. 2* (Ed.: H. van de - Waterbeemd), Wiley-VCH, Weinheim, **1995**.
[155] Y. C. Martin in *Designing Bioactive Molecules* (Eds.: Y. C. Martin, P. Willett, S. R. Heller), Americal Chemical Society, Washington, **1999**.
[156] D. P. Marriot, I. A. Dougall, P. Meghani, Y. Liu, D. R. Flower, *J. Med. Chem.* **1999**, *42*, 3210 – 3216.
[157] J. P. Snyder, S. N. Rao, K. F. Koehler, A. Vedani in *3D QSAR in Drug Design. Theory, Methods and Applications* (Ed.: H. Kubinyi), Escom, Leiden, **1993**, pp. 336 – 354.
[158] A. Vedani, P. Zbinden, J. P. Snyder, P. A. Greenidge, *J. Am. Chem. Soc.* **1995**, *117*, 4987 – 4994.
[159] M. Gurrath, G. Müller, H.-D. Höltje in *3D QSAR in Drug Design. Recent Advances* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/Escom, Dordrecht, **1998**, p. 135 – 157.
[160] A. Schafferhans, G. Klebe, *J. Mol. Biol.* **2001**, *307*, 407 – 427.
[161] M. S. Doscher, F. M. Richards, *J. Biol. Chem.* **1963**, *238*, 2399 – 2406.
[162] C. C. F. Blake, R. Cassels, C. M. Dobson, F. M. Poulsen, R. J. P. Williams, K. S. Wilson, *J. Mol. Biol.* **1981**, *147*, 73 – 95.
[163] A. T. Brünger, *Nat. Struct. Biol.* **1997**, *4 (NMR Suppl.)*, 862 – 865.
[164] A. L. Margolin, *Trends Biotechnol.* **1996**, *14*, 223 – 230.
[165] A. Mozzarelli, G. L. Rossi, *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 343 – 365.
[166] A. L. Margolin, M. A. Navia, *Angew. Chem.* **2001**, *113*, 2262 – 2281; *Angew. Chem. Int. Ed.* **2001**, *40*, 2205 – 2222.
[167] M. A. Murcko, *Rev. Comput. Chem.* **1997**, *11*, 1 – 66.
[168] D. E. Clark, C. W. Murray, J. Li, *Rev. Comput. Chem.* **1997**, *11*, 67 – 125.
[169] T. Lengauer, M. Rarey, *Curr. Opin. Struct. Biol.* **1996**, *6*, 402 – 406.
[170] H. J. Böhm, M. Stahl, *Curr. Opin. Chem. Biol.* **2000**, *4*, 283 – 286.
[171] G. M. Verkhivker, D. Bouzida, D. K. Gehlhaar, P. A. Rejto, S. Arthurs, A. B. Colson, S. T. Freer, V. Larson, B. A. Luty, T. Marrone, P. W. Rose, *J. Comput. Aided Mol. Des.* **2000**, *14*, 731 – 751.
[172] J. S. Dixon, *Proteins* **1997**, *Suppl. 1*, 198 – 204.
[173] Ajay, M. A. Murcko, *J. Med. Chem.* **1995**, *38*, 4953 – 4967.
[174] R. M. A. Knegtel, P. D. J. Grootenhuis in *3D QSAR in Drug Design. Ligand Protein Interactions and Molecular Similarity* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/Escom, Dordrecht, **1998**, pp. 99 – 114.
[175] T. I. Oprea, G. R. Marshall in *3D QSAR in Drug Design. Ligand Protein Interactions and Molecular Similarity* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/Escom, Dordrecht, **1998**, pp. 3 – 17.
[176] B. Kramer, M. Rarey, T. Lengauer, *Proteins* **1999**, *37*, 145 – 156.
[177] H. A. Carlson, J. A. McCammon, *Mol. Pharmacol.* **2000**, *57*, 213 – 218.

[178] N. E. Chayen, T. J. Boggon, A. Casetta, A. Deacon, T. Gleichmann, J. Habash, S. J. Harrop, J. R. Helliwell, Y. P. Nieh, M. R. Peterson, J. Raftery, E. H. Snell, A. Hädener, A. C. Niemann, D. P. Siddons, V. Stojanoff, A. W. Thompson, T. Ursby, M. Wulff, *Q. Rev. Biophys.* **1996**, *29*, 227 – 278.
[179] R. W. W. Hooft, C. Sander, G. Vriend, *Proteins* **1996**, *26*, 363 – 376.
[180] J. E. Nielsen, G. Vriend, *Proteins* **2001**, *43*, 403 – 412.
[181] S. Longhi, M. Czjzek, C. Cambillau, *Curr. Opin. Struct. Biol.* **1998**, *8*, 730 – 737.
[182] R. Flaig, T. Koritsanszky, R. Soyka, L. Häming, P. Luger, *Angew. Chem.* **2001**, *113*, 368 – 371; *Angew. Chem. Int. Ed.* **2001**, *40*, 355 – 359.
[183] A. A. Kossiakoff, M. Randal, J. Guenot, C. Eigenbrot, *Proteins* **1992**, *14*, 65 – 74.
[184] A. Wlodawer, J. Nachman, G. L. Gilliland, W. Gallagher, C. J. Woodward, *J. Mol. Biol.* **1987**, *198*, 469 – 480.
[185] P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, A. T. Hagler, *Proteins* **1988**, *4*, 31 – 47.
[186] M. Levitt, B. H. Park, *Structure* **1993**, *1*, 223 – 226.
[187] O. Carugo, D. Bordo, *Acta Crystallogr. Sect. D* **1999**, *55*, 479 – 483.
[188] P. A. Karplus, C. Faerman, *Curr. Opin. Struct. Biol.* **1994**, *4*, 770 – 776.
[189] T. Beyer, G. M. Day, S. L. Price, *J. Am. Chem. Soc.* **2001**, *123*, 5086 – 5094.
[190] J. Bernstein, *Prog. Clin. Biol. Res.* **1989**, *289*, 203 – 215.
[191] P. Verwer, F. J. J. Leusen, *Rev. Comput. Chem.* **1998**, *12*, 327 – 365.
[192] J. D. Dunitz, J. Bernstein, *Acc. Chem. Res.* **1995**, *28*, 193 – 200.
[193] J. M. Lehn, *Angew. Chem.* **1988**, *100*, 91 – 116; *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 89 – 112.
[194] H.-J. Böhm, G. Klebe, *Angew. Chem.* **1996**, *108*, 2750 – 2778; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2589 – 2614.
[195] G. Klebe, H.-J. Böhm, *Period. Biol.* **1998**, *100 (Suppl. 2)*, 77 – 83.
[196] E. di Cera, *Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules*, Cambridge University Press, Cambridge, **1995**.
[197] *Drug – Receptor Thermodynamics: Introduction and Applications* (Ed.: R. B. Raffa), Wiley, Chichester, **2001**.
[198] M. K. Gilson, J. A. Given, B. L. Bush, J. A. McCammon, *Biophys. J.* **1997**, *72*, 1047 – 1069.
[199] A. M. Davis, S. J. Teague, *Angew. Chem.* **1999**, *111*, 778 – 792; *Angew. Chem. Int. Ed.* **1999**, *38*, 737 – 749.
[200] I. D. Kuntz, K. Chen, K. A. Sharp, P. A. Kollman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9997 – 10002.
[201] A. Fersht, *Enzyme Structure and Mechanism*, Freeman, New York, **1985**.
[202] P. M. Dean, *Molecular Foundations of Drug – Receptor Interaction*, Cambridge University Press, Cambridge, **1987**.
[203] P. R. Andrews, D. J. Craik, J. L. Martin, *J. Med. Chem.* **1984**, *27*, 1648 – 1657.
[204] J. D. Dunitz, A. Gavezotti, *Acc. Chem. Res.* **1999**, *32*, 677 – 684.
[205] G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer, Berlin, **1991**.
[206] G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, New York, **1997**.
[207] L. Pauling, R. B. Corey, *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 729 – 740.
[208] K. P. Murphy, S. J. Gill, *J. Mol. Biol.* **1991**, *222*, 699 – 706.
[209] A. Yang, K. A. Sharp, B. Honig, *J. Mol. Biol.* **1992**, *227*, 889 – 900.
[210] A. R. Fersht, *Trends Biochem. Sci.* **1987**, *12*, 301 – 304.
[211] Q. Zou, S. M. Habermann-Rottinghaus, K. P. Murphy, *Proteins* **1998**, *31*, 107 – 115.
[212] H. Adalsteinsson, A. H. Maulitz, T. C. Bruice, *J. Am. Chem. Soc.* **1996**, *118*, 7689 – 7693.
[213] T. Steiner, W. Saenger, *Acta Crystallogr. Sect. B* **1994**, *50*, 348 – 357.
[214] Z. Wang, H. Luecke, Y. Nanhua, F. Quiocho, *Nat. Struct. Biol.* **1997**, *4*, 519 – 522.
[215] A. Warshel, A. Papazyan, P. A. Kollman, *Science* **1995**, *269*, 102 – 104.
[216] B. Honig, A. Nicholls, *Science* **1995**, *268*, 1144 – 1149.
[217] T. Simonson, G. Archontis, M. Karplus, *J. Phys. Chem. B* **1999**, *103*, 6142 – 6156.
[218] H. Nakamura, *Q. Rev. Biophys.* **1996**, *29*, 1 – 90.
[219] W. Rocchia, E. Alexov, B. Honig, *J. Phys. Chem. B* **2001**, *105*, 6507 – 6514.
[220] M. Stahl, H.-J. Böhm, *J. Mol. Graphics Modell.* **1998**, *16*, 121 – 132.
[221] C. Beeson, N. Pham, G. Shipps, T. A. Dix, *J. Am. Chem. Soc.* **1993**, *115*, 6803 – 6812.
[222] I. K. McDonald, J. M. Thornton, *J. Mol. Biol.* **1994**, *238*, 777 – 793.
[223] P. R. Connelly, R. A. Aldape, F. J. Bruzzese, S. P. Chambers, M. J. Fitzgibbon, M. A. Fleming, S. Itoh, D. J. Livingston, M. A. Navia, J. A. Thomson, K. P. Wilson, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 1964 – 1968.
[224] D. Grobelny, U. B. Goli, R. E. Galardy, *Biochemistry* **1989**, *28*, 4948 – 4951.
[225] P. A. Bash, U. C. Singh, F. K. Brown, R. Langridge, P. A. Kollman, *Science* **1987**, *235*, 574 – 576.
[226] A. R. Fersht, J. S. Shindler, W. C. Tsui, *Biochemistry* **1980**, *19*, 5520 – 5524.
[227] R. M. C. Dawson, D. C. Elliott, W. H. Elliot, K. M. Jones, *Data for Biochemical Research*, Oxford University Press, Oxford, **1969**.
[228] J. Antosiewicz, J. A. McCammon, M. K. Gilson, *Biochemistry* **1996**, *35*, 7819 – 7833.
[229] M. Graffner-Nordberg, J. Marelius, S. Ohlsson, A. Persson, G. Swedberg, P. Andersson, S. E. Andersson, J. Åqvist, A. Hallberg, *J. Med. Chem.* **2000**, *43*, 3852 – 3861.
[230] X. Barril, C. Aleman, M. Orozco, F. J. Luque, *Proteins* **1998**, *32*, 67 – 79.
[231] A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye, G. Winter, *Nature* **1985**, *314*, 235 – 238.
[232] S. M. Habermann, K. P. Murphy, *Protein Sci.* **1996**, *5*, 1229 – 1239.
[233] K. A. Dill, *Biochemistry* **1990**, *29*, 7133 – 7155.
[234] J. S. Thorson, E. Chapman, E. C. Murphy, P. G. Schultz, J. K. Judice, *J. Am. Chem. Soc.* **1995**, *117*, 1157 – 1158.
[235] J. Fernandez-Recio, A. Romero, J. Sancho, *J. Mol. Biol.* **1999**, *290*, 319 – 330.
[236] M. A. Hossain, H.-J. Schneider, *Chem. Eur. J.* **1999**, *5*, 1284 – 1290.
[237] D. E. Otzen, A. R. Fersht, *Protein Eng.* **1999**, *12*, 41 – 45.
[238] J. Gao, K. Kuczera, B. Tidor, M. Karplus, *Science* **1989**, *244*, 1069 – 1072.
[239] A. J. Doig, D. H. Williams, *J. Am. Chem. Soc.* **1992**, *114*, 338 – 343.
[240] M. S. Searle, D. H. Williams, U. Gerhard, *J. Am. Chem. Soc.* **1992**, *114*, 10697 – 10704.
[241] D. H. Williams, M. S. Searle, J. P. Mackay, U. Gerhard, R. A. Maplestone, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1172 – 1178.
[242] D. H. Williams, M. S. Westwell, *Chem. Soc. Rev.* **1998**, *27*, 57 – 63.
[243] M. S. Gordon, J. H. Jensen, *Acc. Chem. Res.* **1996**, *29*, 536 – 543.
[244] G. Klebe, *J. Mol. Biol.* **1994**, *237*, 212 – 235.
[245] F. H. Allen, *Acta Crystallogr. Sect. A* **1998**, *54*, 758 – 771.
[246] J. Kroon, J. A. Kanters, J. G. C. M. van Duijneveldt-van de Rijdt, F. B. van Duijneveldt, J. A. Vliegenhardt, *J. Mol. Struct.* **1975**, *24*, 109 – 129.
[247] P. Murray-Rust, J. P. Glusker, *J. Am. Chem. Soc.* **1984**, *106*, 1018 – 1025.
[248] M. Tintelnot, P. Andrews, *J. Comput. Aided Mol. Des.* **1989**, *3*, 67 – 84.
[249] S. M. Roe, M. M. Teeter, *J. Mol. Biol.* **1993**, *229*, 419 – 427.
[250] I. J. Bruno, J. C. Cole, J. P. Lommerse, R. S. Rowland, R. Taylor, M. L. Verdonk, *J. Comput. Aided Mol. Des.* **1997**, *11*, 525 – 537.
[251] J. C. Cole, R. Taylor, M. L. Verdonk, *Acta Crystallogr. Sect. D* **1998**, *54*, 1183 – 1193.
[252] R. Taylor, O. Kennard, *J. Am. Chem. Soc.* **1982**, *104*, 5063 – 5070.
[253] S. Harder, *Chem. Eur. J.* **1999**, *5*, 1852 – 1861.
[254] G. R. Desiraju, *Acc. Chem. Res.* **1996**, *29*, 441 – 449.
[255] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J. Am. Chem. Soc.* **2000**, *122*, 11450 – 11458.
[256] M. Brandl, M. S. Weiss, A. Jabs, J. Sühnel, R. Hilgenfeld, *J. Mol. Biol.* **2001**, *307*, 357 – 377.
[257] C. A. Hunter, J. K. M. Sanders, *J. Am. Chem. Soc.* **1990**, *112*, 5525 – 5534.
[258] S. K. Burley, G. A. Petsko, *Science* **1985**, *229*, 23 – 28.
[259] U. Samanta, D. Pal, P. Charkrabarti, *Acta Crystallogr. Sect. D* **1999**, *55*, 1421 – 1427.
[260] C. A. Hunter, J. Singh, J. M. Thornton, *J. Mol. Biol.* **1991**, *218*, 837 – 846.
[261] G. B. McGaughey, M. Gagné, A. K. Rappé, *J. Biol. Chem.* **1998**, *273*, 15458 – 15463.

[262] D. A. Dougherty, *Science* **1996**, *271*, 163 – 168.
[263] J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303 – 1324.
[264] E. V. Pletneva, A. T. Laederach, D. B. Fulton, N. M. Kostic, *J. Am. Chem. Soc.* **2001**, *123*, 6232 – 6245.
[265] J. D. Schmitt, *Curr. Med. Chem.* **2000**, *7*, 749 – 800.
[266] K. Brejc, W. J. van Dijk, R. V. Klaasen, M. Schuurmans, J. van der Oost, A. B. Smit, T. K. Sixma, *Nature* **2001**, *411*, 269 – 276.
[267] J. P. Gallivan, D. A. Dougherty, *J. Am. Chem. Soc.* **2000**, *122*, 870 – 874.
[268] M. M. Harding, *Acta Crystallogr. Sect. D* **1999**, *55*, 1432 – 1443.
[269] M. M. Harding, *Acta Crystallogr. Sect. D* **2000**, *56*, 857 – 867.
[270] C.-Y. Kim, J. S. Chang, J. B. Doyon, T. T. Baird, C. A. Fierke, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2000**, *122*, 12125 – 12134.
[271] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373 – 380.
[272] J. Israelachvili, H. Wennerstrom, *Nature* **1996**, *379*, 219 – 225.
[273] D. G. Covell, A. Wallqvist, *J. Mol. Biol.* **1997**, *269*, 281 – 297.
[274] C. S. Poornima, P. M. Dean, *J. Comput. Aided Mol. Des.* **1995**, *9*, 500 – 512.
[275] H. Wang, A. Ben-Naim, *J. Med. Chem.* **1996**, *39*, 1531 – 1539.
[276] A. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists*, Plenum, New York, **1992**.
[277] J. E. Ladbury, *Chem. Biol.* **1996**, *3*, 973 – 980.
[278] J. D. Dunitz, *Science* **1994**, *264*, 670.
[279] *Structural-Based Drug Design: Thermodynamics, Modeling and Strategy* (Ed.: P. R. Connelly), Landes, Austin, **1997**.
[280] C. S. Poornima, P. M. Dean, *J. Comput. Aided Mol. Des.* **1995**, *9*, 513 – 520.
[281] L. A. Kuhn, M. A. Siani, M. E. Pique, C. L. Fisher, E. D. Getzoff, J. A. Trainer, *J. Mol. Biol.* **1992**, *228*, 13 – 22.
[282] J. R. H. Tame, A. J. Wilkinson, *Science* **1994**, *264*, 1578 – 1581.
[283] J. R. H. Tame, S. H. Sleigh, A. J. Wilkinson, J. E. Ladbury, *Nat. Struct. Biol.* **1996**, *3*, 998 – 1001.
[284] C. S. Poornima, P. M. Dean, *J. Comput. Aided Mol. Des.* **1995**, *9*, 521 – 531.
[285] M. L. Raymer, P. C. Sanschagrin, W. F. Punch, S. Venkataraman, E. D. Goodman, L. A. Kuhn, *J. Mol. Biol.* **1997**, *265*, 445 – 464.
[286] D. Ringe, C. Mattos, *Med. Res. Rev.* **1999**, *19*, 321 – 331.
[287] *Water, a Comprehensive Treatise*, *Vol. 1 – 7* (Ed.: F. Franks), Plenum, New York, **1972 – 1982**.
[288] D. Eisenberg, W. Kauzmann, *The Structure and Properties of Water*, Oxford University Press, Oxford, **1969**.
[289] R. Ludwig, *Angew. Chem.* **2001**, *113*, 1856 – 1876; *Angew. Chem. Int. Ed.* **2001**, *40*, 1809 – 1827.
[290] M. Page, *Angew. Chem.* **1977**, *89*, 456 – 467; *Angew. Chem. Int. Ed. Engl.* **1977**, *16*, 449 – 459.
[291] W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046 – 4050.
[292] M. S. Searle, D. H. Williams, *J. Am. Chem. Soc.* **1992**, *114*, 10690 – 10697.
[293] E. Grunwald, C. Steel, *J. Am. Chem. Soc.* **1995**, *117*, 5687 – 5692.
[294] C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, Wiley, New York, **1980**.
[295] A. Ben-Naim, *Hydrophobic Interactions*, Plenum, New York, **1980**.
[296] A. Ben-Naim, *Solvation Thermodynamics*, Plenum, New York, **1987**.
[297] K. A. T. Silverstein, A. D. J. Haymet, K. A. Dill, *J. Am. Chem. Soc.* **1998**, *120*, 3166 – 3175.
[298] W. Blokzijl, J. B. F. N. Engberts, *Angew. Chem.* **1993**, *105*, 1610 – 1648; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1545 – 1579.
[299] H. S. Frank, M. W. Evans, *J. Chem. Phys.* **1945**, *13*, 507.
[300] G. Nemethy, H. A. Scheraga, *J. Chem. Phys.* **1962**, *36*, 3401.
[301] K. A. T. Silverstein, A. D. J. Haymet, K. A. Dill, *J. Am. Chem. Soc.* **2000**, *122*, 8037 – 8041.
[302] N. Muller, *Acc. Chem. Res.* **1990**, *23*, 23 – 28.
[303] A. Pertsemlidis, A. M. Saxena, A. K. Soper, T. Head-Gordon, R. M. Glaeser, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10769.
[304] K. E. Laidig, V. Daggett, *J. Phys. Chem.* **1996**, *100*, 5616 – 5619.
[305] L. F. Scatena, M. G. Brown, G. L. Richmond, *Science* **2001**, *292*, 908 – 912.
[306] R. L. Mancera, *J. Comput. Aided Mol. Des.* **1996**, *10*, 321 – 326.
[307] K. P. Murphy, P. L. Privalov, S. J. Gill, *Science* **1990**, *247*, 559.
[308] M. C. Chervenak, E. J. Toone, *J. Am. Chem. Soc.* **1994**, *116*, 10533 – 10539.
[309] C. Chothia, *Nature* **1974**, *248*, 338 – 339.
[310] C. Chothia, J. Janin, *Nature* **1975**, *256*, 705 – 708.
[311] T. Ooi, M. Oobatake, G. Nemethy, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 3086 – 3090.
[312] D. Eisenberg, A. D. McLachlan, *Nature* **1986**, *319*, 199 – 203.
[313] C. N. Pace, *J. Mol. Biol.* **1992**, *226*, 29 – 35.
[314] J. Wang, W. Wang, S. Huo, M. Lee, P. A. Kollman, *J. Phys. Chem. B* **2001**, *105*, 5055 – 5067.
[315] D. H. Williams, B. Bardsley, *Perspect. Drug Discovery Des.* **1999**, *17*, 43 – 59.
[316] P. Carrupt, B. Testa, P. Gaillard, *Rev. Comput. Chem.* **1997**, *11*, 241 – 315.
[317] J. A. Reynolds, D. B. Gilbert, C. Tanford, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 2925 – 2927.
[318] R. B. Hermann, *J. Phys. Chem.* **1972**, *76*, 2754.
[319] A. Nicholls, K. A. Sharp, B. Honig, *Proteins* **1991**, *11*, 281 – 296.
[320] P. D. Ross, M. V. Rekharsky, *Biophys. J.* **1996**, *71*, 2144 – 2155.
[321] L. Serrano, J.-L. Neira, J. Sancho, A. R. Fersht, *Nature* **1992**, *356*, 453 – 455.
[322] J. T. Kellis, K. Nyberg, A. R. Fersht, *Biochemistry* **1989**, *28*, 4914 – 4922.
[323] D. Shortle, W. E. Stites, A. K. Meeker, *Biochemistry* **1990**, *29*, 8033 – 8041.
[324] M. Matsumura, W. J. Becktel, B. W. Matthews, *Nature* **1988**, *334*, 406 – 410.
[325] G. J. Sharman, M. S. Searle, B. Benhamu, P. Groves, D. H. Williams, *Angew. Chem.* **1995**, *107*, 1644 – 1646; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1483 – 1485.
[326] D. M. Huang, D. Chandler, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8324 – 8327.
[327] M. Vandonselaar, R. A. Hickie, J. W. Quail, L. T. J. Delbaere, *Nat. Struct. Biol.* **1994**, *1*, 795 – 801.
[328] H. P. Märki, W. Fischli, A. Binggeli, V. Breu, D. Bur, J. P. Clozel, A. D'Arcy, F. Grüninger, R. Güller, G. Hirth, T. Lave, S. Mathews, M. Müller, C. Oefner, H. Stadler, E. Vieira, M. Wilhelm, W. Wostl, *9th RSC-SCI Medicinal Chemistry Symposium* (Cambridge, UK), **1997**.
[329] A. Urzhumtsev, F. Tete-Favier, A. Mitschler, J. Barbanton, P. Barth, L. Urzhumtseva, J. F. Biellmann, A. Podjarny, D. Moras, *Structure* **1997**, *5*, 601 – 612.
[330] S. E. Zographos, N. G. Oikonomakos, K. E. Tsitsanou, D. D. Leonidas, E. D. Chrysina, V. T. Skamnaki, H. Bischoff, S. Goldmann, K. A. Watson, L. N. Johnson, *Structure* **1997**, *5*, 1413 – 1425.
[331] J. Ren, R. Esnouf, E. Garman, D. Somers, C. Ross, I. Kirby, J. Keeling, G. Dardy, Y. Jones, D. Stuart, D. Stammers, *Nat. Struct. Biol.* **1995**, *2*, 293 – 302.
[332] G. P. Brady, K. A. Sharp, *Curr. Opin. Struct. Biol.* **1997**, *7*, 215 – 221.
[333] M. I. Page, W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678 – 1683.
[334] M. I. Page, *Chem. Soc. Rev.* **1973**, *2*, 295 – 323.
[335] J. M. Sturtevant, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 2236 – 2240.
[336] J. D. Dunitz, *Chem. Biol.* **1995**, *2*, 709 – 712.
[337] P. W. Atkins, *Physikalische Chemie*, VCH, Weinheim, **1990**.
[338] K. P. Murphy, D. Xie, K. Thompson, M. Arnzel, E. Freire, *Proteins* **1994**, *18*, 63 – 67.
[339] J. Janin, *Proteins* **1996**, *24*, R1 – R2.
[340] J. Janin, *Proteins* **1995**, *21*, 30 – 39.
[341] J. Doucet, J. P. Benoit, *Nature* **1987**, *325*, 643.
[342] D. L. D. Caspar, J. Clarage, D. M. Salunke, M. Clarage, *Nature* **1988**, *332*, 659 – 662.
[343] A. V. Finkelstein, J. Janin, *Protein Eng.* **1989**, *3*, 1 – 3.
[344] G. P. Brady, K. A. Sharp, *Biophys. J.* **1997**, *72*, 913 – 927.
[345] J. Hermans, L. Wang, *J. Am. Chem. Soc.* **1997**, *119*, 2707 – 2714.
[346] S. D. Pickett, M. J. Sternberg, *J. Mol. Biol.* **1993**, *231*, 825 – 839.
[347] A. J. Doig, M. J. E. Sternberg, *Protein Sci.* **1995**, *4*, 2247 – 2251.
[348] J. D. Forman-Kay, *Nat. Struct. Biol.* **1999**, *6*, 1086 – 1087.
[349] R. Raag, T. L. Poulos, *Biochemistry* **1991**, *30*, 2674 – 2684.
[350] L. Yu, C. X. Zhu, Y. C. Tse-Dinh, S. W. Fesik, *Biochemistry* **1996**, *35*, 9661 – 9666.
[351] L. Zidek, M. V. Novotny, M. J. Stone, *Nat. Struct. Biol.* **1999**, *6*, 1118 – 1121.
[352] P. C. Weber, J. J. Wendoloski, M. W. Pantoliano, F. R. Salemme, *J. Am. Chem. Soc.* **1992**, *114*, 3197 – 3200.
[353] M. S. L. Lim, E. R. Johnston, C. A. Kettner, *J. Med. Chem.* **1993**, *36*, 1831 – 1838.

[354] R. A. Wiley, D. H. Rich, *Med. Res. Rev.* **1993**, *13*, 327 – 384.
[355] B. Testa, P. A. Carrupt, P. Gaillard, F. Billois, P. Weber, *Pharm. Res.* **1996**, *13*, 335 – 343.
[356] M. A. Navia, P. R. Chaturvedi, *Drug Discovery Today* **1996**, *1*, 179 – 189.
[357] M. C. Nicklaus, S. Wang, J. S. Driscoll, G. W. A. Milne, *Bioorg. Med. Chem.* **1995**, *3*, 411 – 428.
[358] M. J. Spark, D. A. Winkler, P. R. Andrews, *Int. J. Quantum Chem.* **1982**, *9*, 321.
[359] J. Bostrom, P. O. Norrby, T. Liljefors, *J. Comput. Aided Mol. Des.* **1998**, *12*, 383 – 396.
[360] M. Vieth, J. D. Hirst, C. L. Brooks III, *J. Comput. Aided Mol. Des.* **1998**, *12*, 563 – 572.
[361] D. Mastropaolo, A. Camerman, N. Camerman, *J. Med. Chem.* **2001**, *44*, 269 – 273.
[362] T. A. Halgren, W. Damm, *Curr. Opin. Struct. Biol.* **2001**, *11*, 236 – 242.
[363] W. F. Lau, B. M. Pettitt, *J. Med. Chem.* **1989**, *32*, 2542 – 2547.
[364] A. E. Mark, W. F. van Gunsteren, *J. Mol. Biol.* **1994**, *240*, 167 – 176.
[365] A. Ben-Naim, *J. Chem. Phys.* **1997**, *107*, 3698 – 3706.
[366] G. F. Ackers, F. R. Smith, *Annu. Rev. Biochem.* **1985**, *54*, 597 – 629.
[367] A. Horovitz, *J. Mol. Biol.* **1987**, *196*, 733 – 735.
[368] J. A. Wells, *Biochemistry* **1990**, *29*, 8509 – 8515.
[369] K. A. Dill, *J. Biol. Chem.* **1997**, *272*, 701 – 704.
[370] G. P. Brady, K. A. Sharp, *J. Mol. Biol.* **1995**, *254*, 77 – 85.
[371] G. K. Ackers, M. L. Doyle, D. Myers, M. A. Daugherty, *Science* **1992**, *255*, 54 – 63.
[372] D. H. Williams, A. J. Maguire, W. Tsuzuki, M. S. Westwell, *Science* **1998**, *280*, 711 – 714.
[373] T. Wieseman, S. Wiliston, J. Brandts, L. Lin, *Anal. Biochem.* **1989**, *179*, 131 – 137.
[374] R. Hitzemann, *Trends Pharmacol. Sci.* **1988**, *9*, 408.
[375] A. F. D. de Namor, M.-C. Ritt, M.-J. Schwing-Weill, F. Arnaud-Neu, D. F. V. Lewis, *J. Chem. Soc. Faraday Trans.* **1991**, *87*, 3231 – 3239.
[376] P. Gilli, V. Ferretti, G. Gilli, P. A. Borea, *J. Phys. Chem.* **1994**, *98*, 1515 – 1518.
[377] C. T. Calderone, D. H. Williams, *J. Am. Chem. Soc.* **2001**, *123*, 6262 – 6267.
[378] E. Gallicchio, M. M. Kubo, R. M. Levy, *J. Am. Chem. Soc.* **1998**, *120*, 4526 – 4527.
[379] K. Sharp, *Protein Sci.* **2001**, *10*, 661 – 667.
[380] G. Klebe in *Rational Molecular Design in Drug Research* (Eds.: T. Liljefors, F. S. Jorgensen, P. Krogsgaard-Larsen), Munksgaard, Kopenhagen, **1998**, pp. 151 – 160.
[381] H. Kubinyi, *Drug Discovery Today* **1997**, *2*, 457 – 467.
[382] K. H. Kim, G. Greco, E. Novellino, *Perspect. Drug Discovery Des.* **1998**, *12*, 257 – 315.
[383] M. Gurrath, *Curr. Med. Chem.* **2001**, *8*, 1605 – 1648.
[384] S. Halazy, *Expert Opin. Ther. Pat.* **1999**, *9*, 431 – 446.
[385] B. Fontaine, E. Plassart-Schiess, S. Nicole, *Mol. Aspects Med.* **1997**, *18*, 415 – 463.
[386] S. J. Oh, H. J. Ha, D. Y. Chi, H. K. Lee, *Curr. Med. Chem.* **2001**, *8*, 999 – 1034.
[387] P. Schloss, D. C. Williams, *J. Psychopharmacol.* **1998**, *12*, 115 – 121.
[388] R. J. Vandenberg, *Clin. Exp. Pharmacol. Physiol.* **1998**, *25*, 393 – 400.
[389] J. A. McCammon, *Curr. Opin. Struct. Biol.* **1998**, *8*, 245 – 249.
[390] J. R. H. Tame, *J. Comput. Aided Mol. Des.* **1999**, *13*, 99 – 108.
[391] M. R. Reddy, V. N. Viswanadhan, M. D. Erion in *3D QSAR in Drug Design: Ligand Protein Interactions and Molecular Similarity* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/Escom, Leiden, **1998**, pp. 85 – 98.
[392] J. D. Hirst, *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 28 – 33.
[393] P. A. Kollman, *Curr. Opin. Struct. Biol.* **1994**, *4*, 240 – 245.
[394] H. Gohlke, G. Klebe, *Curr. Opin. Struct. Biol.* **2001**, *11*, 231 – 235.
[395] R. D. Brown, Y. C. Martin, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572 – 584.
[396] C. Humblet, J. B. Dunbar, *Annu. Rep. Med. Chem.* **1993**, *28*, 275 – 284.
[397] P. Willett, *J. Mol. Recognit.* **1995**, *8*, 290 – 303.
[398] G. Klebe in *3D QSAR in Drug Design. Theory, Methods and Applications* (Ed.: H. Kubinyi), Escom, Leiden, **1993**, pp. 173 – 199.
[399] M. G. Bures in *Practical Application of Computer-Aided Drug Design* (Ed.: P. S. Charifson), Marcel Dekker, New York, **1997**, pp. 39 – 72.
[400] C. Lemmen, T. Lengauer, *J. Comput. Aided Mol. Des.* **2000**, *14*, 215 – 232.
[401] C. Hansch, T. Fujita, *J. Am. Chem. Soc.* **1964**, *86*, 1616 – 1626.
[402] S. M. Free, J. W. Wilson, *J. Med. Chem.* **1964**, *7*, 395 – 399.
[403] K. H. Kim in *3D QSAR in Drug Design. Theory, Methods and Applications* (Ed.: H. Kubinyi), Escom, Leiden, **1993**, pp. 619 – 642.
[404] G. Greco, E. Novellino, Y. C. Martin in *Designing Bioactive Molecules* (Eds.: Y. C. Martin, P. Willett, S. R. Heller), American Chemical Society, Washington, **1998**, pp. 219 – 251.
[405] R. D. Cramer III, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc.* **1988**, *110*, 5959.
[406] C. L. Waller, T. I. Oprea, A. Giolitti, G. R. Marshall, *J. Med. Chem.* **1993**, *36*, 4152.
[407] A. Golbraikh, P. Bernard, J. R. Chrétien, *Eur. J. Med. Chem.* **2000**, *35*, 123 – 136.
[408] G. R. Marshall, C. D. Barry, H. E. Bosshard, R. A. Dammkoehler, D. A. Dunn in *Computer-Assisted Drug Design* (Eds.: E. C. Olson, R. E. Christoffersen), American Chemical Society, Washington, **1979**, pp. 205 – 226.
[409] Y. C. Martin, M. G. Bures, E. A. Danaher, J. DeLazzer in *Trends in QSAR and Molecular Modelling 92* (Ed.: C.-G. Wermuth), Escom, Leiden, **1993**, pp. 20 – 27.
[410] *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications* (Eds.: F. Sanz, J. Giraldo, F. Manaut), Prous, Barcelona, **1995**.
[411] *Chemometric Methods in Molecular Design*, *Vol. 2* (Ed.: H. van de Waterbeemd), VCH, Weinheim, **1995**.
[412] R. D. Cramer III, J. D. Bunce in *QSAR in Drug Design and Toxicology* (Eds.: D. Hadzi, B. Jerman-Blasiz), Elsevier, Amsterdam, **1987**.
[413] P. Geladi, B. R. Kowalski, *Anal. Chim. Acta* **1986**, *185*, 1 – 17.
[414] S. Wold, E. Johansson, M. Cocchi in *3D QSAR in Drug Design. Theory, Methods and Applications* (Ed.: H. Kubinyi), Escom, Leiden, **1993**.
[415] G. Folkers, A. Merz in *Lipophilicity in Drug Action and Toxicology*, *Vol. 4* (Eds.: V. Pliska, B. Testa, H. van de Waterbeemd), VCH, Weinheim, **1996**, pp. 219 – 232.
[416] G. E. Kellogg, D. J. Abraham, *J. Comput. Aided Mol. Des.* **1991**, *5*, 545.
[417] G. E. Kellogg, J. C. Burnett, D. J. Abraham, *J. Comput. Aided Mol. Des.* **2001**, *15*, 381 – 393.
[418] P. J. Goodford, *J. Med. Chem.* **1985**, *28*, 849 – 857.
[419] DelPhi/Solvation, Molecular Simulations Inc., San Diego, CA, **1995**.
[420] G. Klebe, U. Abraham, T. Mietzner, *J. Med. Chem.* **1994**, *37*, 4130 – 4146.
[421] G. Klebe, U. Abraham, *J. Comput. Aided Mol. Des.* **1999**, *13*, 1 – 10.
[422] A. Doweyko, *J. Med. Chem.* **1988**, *31*, 1396 – 1406.
[423] A. N. Jain, K. Koile, D. Chapman, *J. Med. Chem.* **1994**, *37*, 2315 – 2327.
[424] V. E. Golender, E. R. Vorpagel in *3D QSAR in Drug Design. Theory, Methods and Applications* (Ed.: H. Kubinyi), Escom, Leiden, **1993**, pp. 137 – 149.
[425] H.-D. Höltje, L. B. Kier, *J. Pharm. Sci.* **1974**, *63*, 1722 – 1725.
[426] D. Joseph-McCarthy, *Pharm. Ther.* **1999**, *84*, 179 – 191.
[427] H.-J. Böhm, M. Stahl, *Med. Chem. Res.* **1999**, *9*, 445 – 462.
[428] P. Bamborough, F. E. Cohen, *Curr. Opin. Struct. Biol.* **1996**, *6*, 236 – 241.
[429] R. C. Wade in *3D QSAR in Drug Design*, *Vol. 2* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer, Dordrecht, **1998**, pp. 486 – 505.
[430] I. T. Weber, R. W. Harrison in *3D QSAR in Drug Design*, *Vol. 2* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer, Dordrecht, **1998**, 115 – 127.
[431] T. Liljefors in *3D QSAR in Drug Design*, *Vol. 2* (Eds.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer, Dordrecht, **1998**, 3 – 17.
[432] A. Warshel, A. Papazyan, *Curr. Opin. Struct. Biol.* **1998**, *8*, 211 – 217.
[433] M. K. Gilson, *Curr. Opin. Struct. Biol.* **1995**, *5*, 216 – 223.
[434] D. Bashford, D. A. Case, *Annu. Rev. Phys. Chem.* **2000**, *51*, 129 – 152.
[435] M. Orozco, F. J. Luque, *Chem. Rev.* **2000**, *100*, 4187 – 4225.
[436] M. Schaefer, H. W. T. van Vlijmen, M. Karplus, *Adv. Protein Chem.* **1998**, *51*, 1 – 57.

[437] P. A. Kollman, *Acc. Chem. Res.* **1996**, *29*, 461 – 469.
[438] P. Kollman, *Chem. Rev.* **1993**, *93*, 2395 – 2417.
[439] P. A. Kollman, K. M. Merz, Jr., *Acc. Chem. Res.* **1990**, *23*, 246 – 252.
[440] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham III, *Acc. Chem. Res.* **2000**, *33*, 889 – 897.
[441] W. P. Reinhardt, M. A. Miller, L. M. Amon, *Acc. Chem. Res.* **2001**, *34*, 607 – 614.
[442] M. R. Reddy, M. D. Erion, A. Agarwal, *Rev. Comput. Chem.* **2000**, *16*, 217 – 304.
[443] W. L. Jorgensen, *Acc. Chem. Res.* **1989**, *22*, 184 – 189.
[444] T. P. Straatsma, *Rev. Comput. Chem.* **1996**, *9*, 81 – 127.
[445] H. Meirovitch, *Rev. Comput. Chem.* **1998**, *12*, 1 – 74.
[446] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300 – 313.
[447] R. J. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420 – 1426.
[448] J. P. M. Postma, H. J. C. Berendsen, J. R. Haak, *Faraday Symp. Chem. Soc.* **1981**, *17*, 55 – 67.
[449] B. L. Tembe, J. A. McCammon, *Comput. Chem.* **1984**, *4*, 281 – 283.
[450] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **1953**, *21*, 1087 – 1092.
[451] W. F. van Gunsteren, H. J. C. Berendsen, *Angew. Chem.* **1990**, *102*, 1020 – 1055; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992 – 1023.
[452] S. Boresch, M. Karplus, *J. Mol. Biol.* **1995**, *254*, 801 – 807.
[453] G. Archontis, T. Simonson, M. Karplus, *J. Mol. Biol.* **2001**, *306*, 307 – 327.
[454] D. L. Beveridge, F. M. DiCapua, *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431 – 492.
[455] D. A. Pearlman, *J. Phys. Chem.* **1994**, *98*, 1487 – 1493.
[456] S. Sen, L. Nilsson, *J. Comput. Chem.* **1999**, *20*, 877 – 885.
[457] M. A. McCarrick, P. A. Kollman, *J. Comput. Aided Mol. Des.* **1999**, *13*, 109 – 121.
[458] M. L. Plount Price, W. L. Jorgensen, *J. Am. Chem. Soc.* **2000**, *122*, 9455 – 9466.
[459] W. L. Jorgensen, C. Ravimohan, *J. Chem. Phys.* **1985**, *83*, 3050 – 3054.
[460] C. F. Wong, J. A. McCammon, *J. Am. Chem. Soc.* **1986**, *108*, 3830 – 3831.
[461] K. M. Merz, P. A. Kollman, *J. Am. Chem. Soc.* **1989**, *111*, 5649 – 5658.
[462] N. Ota, A. T. Brunger, *Theor. Chem. Acc.* **1997**, *98*, 407 – 435.
[463] N. Ota, C. Stroupe, J. M. S. Ferreira-da-Silva, S. A. Shah, M. Mares-Guia, A. T. Brunger, *Proteins* **1999**, *37*, 641 – 653.
[464] P. R. Gerber, A. E. Mark, W. F. van Gunsteren, *J. Comput. Aided Mol. Des.* **1993**, *7*, 305 – 323.
[465] B. C. Oostenbrink, J. W. Pitera, M. M. H. van Lipzig, J. H. N. Meerman, W. F. van Gunsteren, *J. Med. Chem.* **2000**, *43*, 4594 – 4605.
[466] H. Liu, A. E. Mark, W. F. van Gunsteren, *J. Phys. Chem.* **1996**, *100*, 9485 – 9494.
[467] Z. Guo, C. L. Brooks, X. Kong, *J. Phys. Chem. B* **1998**, *102*, 2032 – 2036.
[468] S. Banba, Z. Guo, C. L. Brooks III, *J. Phys. Chem. B* **2000**, *104*, 6903 – 6910.
[469] J. Åquist, C. Medina, J. E. Samuelsson, *Protein Eng.* **1994**, *7*, 385 – 391.
[470] T. Hansson, J. Marelius, J. Åquist, *J. Comput. Aided Mol. Des.* **1998**, *12*, 27 – 35.
[471] J. Wang, R. Dixon, P. A. Kollman, *Proteins* **1999**, *34*, 69 – 81.
[472] H. A. Carlson, W. L. Jorgensen, *J. Phys. Chem.* **1995**, *99*, 10667 – 10673.
[473] I. D. Wall, A. R. Leach, D. W. Salt, M. G. Ford, J. W. Essex, *J. Med. Chem.* **1999**, *42*, 5142 – 5152.
[474] A. C. Pierce, W. L. Jorgensen, *J. Med. Chem.* **2001**, *44*, 1043 – 1050.
[475] R. C. Rizzo, J. Tirado-Rives, W. L. Jorgensen, *J. Med. Chem.* **2001**, *44*, 145 – 154.
[476] P. D. J. Grootenhuis, P. J. M. van Galen, *Acta Crystallogr. Sect. D* **1995**, *51*, 560 – 566.
[477] P. D. J. Grootenhuis, S. P. van Helden in *Computational Approaches in Supramolecular Chemistry* (Ed.: G. Wipff), Kluwer, Dordrecht, **1994**, pp. 137 – 149.
[478] M. K. Holloway, J. M. Wai, T. A. Halgren, P. M. Fitzgerald, J. P. Vacca, B. D. Dorsey, R. B. Levin, W. J. Thompson, L. J. Chen, S. J. deSolms, N. Gaffin, A. K. Ghosh, E. A. Giuliani, S. L. Graham, J. P. Guare, R. W. Hungate, T. A. Lyle, W. M. Sanders, T. J. Tucker, M. Wiggins, C. M. Wiscount, O. W. Woltersdorf, S. D. Young, P. L. Darke, J. A. Zugay, *J. Med. Chem.* **1995**, *38*, 305 – 317.
[479] D. Joseph-McCarthy, J. M. Hogle, M. Karplus, *Proteins* **1997**, *29*, 32 – 58.
[480] I. V. Kurinov, R. W. Harrison, *Nat. Struct. Biol.* **1994**, *1*, 735 – 743.
[481] A. Checa, A. R. Ortiz, B. de Pascual-Teresa, F. Gago, *J. Med. Chem.* **1997**, *40*, 4136 – 4145.
[482] S. Vajda, Z. Weng, R. Rosenfeld, C. DeLisi, *Biochemistry* **1994**, *33*, 13977 – 13988.
[483] Z. Weng, S. Vajda, C. DeLisi, *Protein Sci.* **1996**, *5*, 614 – 626.
[484] D. H. Williams, J. P. L. Cox, A. J. Doig, M. Gardener, U. Gerhard, P. T. Kaye, A. R. Lal, I. A. Nicholls, C. J. Salter, R. C. Mitchell, *J. Am. Chem. Soc.* **1991**, *113*, 7020 – 7030.
[485] S. Krystek, T. Stouch, J. Novotny, *J. Mol. Biol.* **1993**, *234*, 661 – 679.
[486] J. Novotny, R. E. Bruccoleri, F. A. Saul, *Biochemistry* **1989**, *28*, 4735 – 4749.
[487] J. Warwicker, H. C. Watson, *J. Mol. Biol.* **1982**, *157*, 671 – 679.
[488] R. J. Zauhar, R. S. Morgan, *J. Mol. Biol.* **1985**, *186*, 815 – 820.
[489] N. Froloff, A. Windemuth, B. Honig, *Protein Sci.* **1997**, *6*, 1293 – 1301.
[490] T. Zhang, D. E. Koshland, *Protein Sci.* **1996**, *5*, 348 – 356.
[491] D. Hoffmann, B. Kramer, T. Washio, T. Steinmetzer, M. Rarey, T. Lengauer, *J. Med. Chem.* **1999**, *42*, 4422 – 4433.
[492] F. Polticelli, P. Ascenzi, M. Bolognesi, B. Honig, *Protein Sci.* **1999**, *8*, 2621 – 2629.
[493] B. K. Shoichet, A. R. Leach, I. D. Kuntz, *Proteins* **1999**, *34*, 4 – 16.
[494] X. Zou, Y. Sun, I. D. Kuntz, *J. Am. Chem. Soc.* **1999**, *121*, 8033 – 8043.
[495] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127 – 6129.
[496] I. Massova, P. A. Kollman, *Perspect. Drug Discovery Des.* **2000**, *18*, 113 – 135.
[497] J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman, *J. Am. Chem. Soc.* **1998**, *120*, 9401 – 9409.
[498] Y. N. Vorobjev, J. C. Almagro, J. Hermans, *Proteins* **1998**, *32*, 399 – 413.
[499] B. Jayaram, D. Sprous, M. A. Young, D. L. Beveridge, *J. Am. Chem. Soc.* **1998**, *120*, 10629 – 10633.
[500] G. Archontis, T. Simonson, M. Karplus, *J. Mol. Biol.* **2001**, *306*, 307 – 327.
[501] C. M. Reyes, P. A. Kollman, *J. Mol. Biol.* **2000**, *297*, 1145 – 1158.
[502] B. Kuhn, P. A. Kollman, *J. Am. Chem. Soc.* **2000**, *122*, 3909 – 3916.
[503] T. Lee, P. A. Kollman, *J. Am. Chem. Soc.* **2000**, *122*, 4385 – 4393.
[504] J. Wang, P. Morin, W. Wang, P. A. Kollman, *J. Am. Chem. Soc.* **2001**, *123*, 5221 – 5230.
[505] T. Lazaridis, M. Karplus, *Proteins* **1999**, *35*, 133 – 152.
[506] J. Gasteiger, J. Zupan, *Angew. Chem.* **1993**, *105*, 510 – 536; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 503 – 527.
[507] H. J. Böhm, *J. Comput. Aided Mol. Des.* **1994**, *8*, 243 – 256.
[508] H. J. Böhm, *J. Comput. Aided Mol. Des.* **1998**, *12*, 309 – 323.
[509] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, R. P. Mee, *J. Comput. Aided Mol. Des.* **1997**, *11*, 425 – 445.
[510] R. Wang, L. Liu, L. Lai, Y. Tang, *J. Mol. Model.* **1998**, *4*, 379 – 394.
[511] C. W. Murray, T. R. Auton, M. D. Eldridge, *J. Comput. Aided Mol. Des.* **1998**, *12*, 503 – 519.
[512] R. D. Head, M. L. Smythe, T. I. Oprea, C. L. Waller, S. M. Green, G. R. Marshall, *J. Am. Chem. Soc.* **1996**, *118*, 3959 – 3969.
[513] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, *J. Am. Chem. Soc.* **1984**, *106*, 765 – 784.
[514] C. Hansch, A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, **1979**.
[515] Y. Takamatsu, A. Itai, *Proteins* **1998**, *33*, 62 – 73.
[516] P. Venkatarangan, A. J. Hopfinger, *J. Med. Chem.* **1999**, *42*, 2169 – 2179.
[517] V. N. Viswanadhan, M. R. Reddy, A. Wlodawer, M. D. Varney, J. N. Weinstein, *J. Med. Chem.* **1996**, *39*, 705 – 712.
[518] D. Rognan, S. L. Lauemoller, A. Holm, S. Buus, V. Tschinke, *J. Med. Chem.* **1999**, *42*, 4650 – 4658.
[519] R. S. Bohacek, C. McMartin, *J. Am. Chem. Soc.* **1994**, *116*, 5560 – 5571.
[520] P. Kasper, P. Christen, H. Gehring, *Proteins* **2000**, *40*, 185 – 192.
[521] A. N. Jain, *J. Comput. Aided Mol. Des.* **1996**, *10*, 427 – 440.

[522] W. Bibel, S. Hölldobler, T. Schaub, *Wissensrepräsentation und Inferenz: Eine grundlegende Einführung*, Vieweg, Braunschweig, **1993**.
[523] R. L. Jernigan, I. Bahar, *Curr. Opin. Struct. Biol.* **1996**, *6*, 195 – 209.
[524] A. E. Torda, *Curr. Opin. Struct. Biol.* **1997**, *7*, 200 – 205.
[525] S. Vajda, M. Sippl, J. Novotny, *Curr. Opin. Struct. Biol.* **1997**, *7*, 222 – 228.
[526] M. J. Sippl, *Curr. Opin. Struct. Biol.* **1995**, *5*, 229 – 235.
[527] H. B. Bürgi, J. D. Dunitz, *Acta Crystallogr. Sect. B* **1988**, *44*, 445 – 448.
[528] P. D. Thomas, K. A. Dill, *J. Mol. Biol.* **1996**, *257*, 457 – 469.
[529] A. V. Finkelstein, A. M. Gutin, A. Y. Badretdinov, *Proteins* **1995**, *23*, 151 – 162.
[530] W. A. Koppensteiner, M. J. Sippl, *Biochemistry* **1998**, *63*, 247 – 252.
[531] T. Lazaridis, M. Karplus, *J. Mol. Biol.* **1999**, *288*, 477 – 487.
[532] J. Moult, *Curr. Opin. Struct. Biol.* **1997**, *7*, 194 – 199.
[533] A. V. Finkelstein, *Curr. Opin. Struct. Biol.* **1997**, *7*, 60 – 71.
[534] M. J. E. Sternberg, P. A. Bates, L. A. Kelley, R. M. MacCallum, *Curr. Opin. Struct. Biol.* **1999**, *9*, 368 – 373.
[535] G. Verkhivker, K. Appelt, S. T. Freer, J. E. Villafranca, *Protein Eng.* **1995**, *8*, 677 – 691.
[536] K. A. Sharp, A. Nicholls, R. Friedman, B. Honig, *Biochemistry* **1991**, *30*, 9686 – 9697.
[537] A. Wallqvist, R. L. Jernigan, D. G. Covell, *Protein Sci.* **1995**, *4*, 1881 – 1903.
[538] R. S. DeWitte, E. I. Shakhnovich, *J. Am. Chem. Soc.* **1996**, *118*, 11733 – 11744.
[539] I. Muegge, Y. C. Martin, *J. Med. Chem.* **1999**, *42*, 791 – 804.
[540] I. Muegge, *J. Comput. Chem.* **2001**, *22*, 418 – 425.
[541] J. B. O. Mitchell, R. A. Laskowski, A. Alex, J. M. Thornton, *J. Comput. Chem.* **1999**, *20*, 1165 – 1176.
[542] K.-C. Ng, W. J. Meath, A. R. Allnatt, *Mol. Phys.* **1979**, *37*, 237 – 253.
[543] J. B. O. Mitchell, R. A. Laskowski, A. Alex, M. J. Forster, J. M. Thornton, *J. Comput. Chem.* **1999**, *20*, 1177 – 1185.
[544] H. Gohlke, M. Hendlich, G. Klebe, *J. Mol. Biol.* **2000**, *295*, 337 – 356.
[545] M. Hendlich, *Acta Crystallogr. Sect. D* **1998**, *54*, 1178 – 1182.
[546] H. Gohlke, M. Hendlich, G. Klebe, *Perspect. Drug Discovery Des.* **2000**, *20*, 115 – 144.
[547] P. S. Charifson, J. J. Corkerey, M. A. Murcko, W. P. Walters, *J. Med. Chem.* **1999**, *42*, 5100 – 5109.
[548] E. C. Meng, B. K. Shoichet, I. D. Kuntz, *J. Comput. Chem.* **1992**, *13*, 505 – 524.
[549] D. K. Gehlhaar, G. M. Verkhivker, P. A. Rejto, C. J. Sherman, D. B. Fogel, S. T. Free, *Chem. Biol.* **1995**, *2*, 317 – 324.
[550] S.-S. So, M. Karplus, *J. Comput. Aided Mol. Des.* **1999**, *13*, 243 – 258.
[551] M. Stahl, M. Rarey, *J. Med. Chem.* **2001**, *44*, 1035 – 1042.
[552] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, *J. Mol. Biol.* **1996**, *261*, 470 – 489.
[553] G. E. Terp, B. N. Johansen, I. T. Christensen, F. S. Jorgensen, *J. Med. Chem.* **2001**, *44*, 2333 – 2343.
[554] D. N. A. Boobbyer, P. J. Goodford, P. M. McWhinnie, R. C. Wade, *J. Med. Chem.* **1989**, *32*, 1083 – 1094.
[555] R. A. Laskowski, J. M. Thornton, C. Humblet, J. Singh, *J. Mol. Biol.* **1996**, *259*, 175 – 201.
[556] M. L. Verdonk, J. C. Cole, R. Taylor, *J. Mol. Biol.* **1999**, *289*, 1093 – 1108.
[557] M. L. Verdonk, J. C. Cole, P. Watson, V. Gillet, P. Willett, *J. Mol. Biol.* **2001**, *307*, 841 – 859.
[558] A. Miranker, M. Karplus, *Proteins* **1991**, *11*, 29 – 34.
[559] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586 – 3616.
[560] C. M. Stultz, M. Karplus, *Proteins* **1999**, *37*, 512 – 529.
[561] R. J. Radmer, P. A. Kollman, *J. Comput. Aided Mol. Des.* **1998**, *12*, 215 – 227.
[562] D. A. Pearlman, *J. Med. Chem.* **1999**, *42*, 4313 – 4324.
[563] C. Bissantz, G. Folkers, D. Rognan, *J. Med. Chem.* **2000**, *43*, 4759 – 4767.
[564] M. R. Reddy, R. J. Bacquet, D. Zichi, D. A. Matthews, K. M. Welsh, T. R. Jones, S. Freer, *J. Am. Chem. Soc.* **1992**, *114*, 10117 – 10122.
[565] M. Rarey, B. Kramer, T. Lengauer in *Proceedings of the 3rd International Conference on Intelligent Systems for Molecular Biology* (Cambridge, UK, **1995**), pp. 300 – 308.
[566] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, T. E. Ferrin, *J. Mol. Biol.* **1982**, *161*, 269 – 288.
[567] H. J. Böhm, *J. Comput. Aided Mol. Des.* **1992**, *6*, 61 – 78.
[568] P. Tao, L. Lai, *J. Comput. Aided Mol. Des.* **2001**, *15*, 429 – 446.
[569] A. Wallqvist, D. G. Covell, *Proteins* **1996**, *25*, 403 – 419.
[570] I. Muegge, Y. C. Martin, P. J. Hajduk, S. W. Fesik, *J. Med. Chem.* **1999**, *42*, 2498 – 2503.
[571] H. Sookhee, R. Andreani, A. Robbins, I. Muegge, *J. Comput. Aided Mol. Des.* **2000**, *14*, 435 – 448.
[572] I. Nobeli, J. B. O. Mitchell, A. Alex, J. M. Thornton, *J. Comput. Chem.* **2001**, *22*, 673 – 688.
[573] R. Hitzemann, *Trends Pharmacol. Sci.* **1988**, *9*, 408 – 411.
[574] H. Roos, R. Karlsson, H. Nilshans, A. Persson, *J. Mol. Recognit.* **1998**, *11*, 204 – 210.
[575] J. Piehler, A. Brecht, G. Gauglitz, C. Maul, S. Grabley, M. Zerlin, *Biosens. Bioelectron.* **1997**, *12*, 531 – 538.
[576] R. L. Rich, D. G. Myszka, *Curr. Opin. Biotechnol.* **2000**, *11*, 54 – 61.
[577] R. P. Hicks, *Curr. Med. Chem.* **2001**, *8*, 627 – 650.
[578] J. M. Moore, *Biopolymers* **1999**, *51*, 221 – 243.
[579] B. Meyer, T. Weimar, T. Paters, *Eur. J. Biochem.* **1997**, *246*, 705 – 709.
[580] W. Jahnke, L. B. Perez, C. G. Paris, A. Strauss, G. Fendrich, C. M. Nalin, *J. Am. Chem. Soc.* **2000**, *122*, 7394 – 7395.
[581] W. Jahnke, S. Rüdisser, M. Zurini, *J. Am. Chem. Soc.* **2001**, *123*, 3149 – 3150.
[582] M. Mayer, B. Meyer, *Angew. Chem.* **1999**, *111*, 1824 – 1844; *Angew. Chem. Int. Ed.* **1999**, *38*, 1784 – 1788.
[583] T. D. Veenstra, *Biophys. Chem.* **1999**, *79*, 63 – 79.
[584] A. A. Rostom, J. H. R. Tame, J. E. Ladbury, C. V. Robinson, *J. Mol. Biol.* **2000**, *296*, 269 – 279.
[585] A. Janshoff, M. Neitzert, Y. Oberdörfer, H. Fuchs, *Angew. Chem.* **2000**, *112*, 3346 – 3374; *Angew. Chem. Int. Ed.* **2000**, *39*, 3212 – 3237.
[586] I. Jelesarov, H. R. Bosshard, *Biochemistry* **1994**, *33*, 13321 – 13328.
[587] R. G. Khalifah, F. Zhang, J. S. Parr, E. S. Rowe, *Biochemistry* **1993**, *32*, 3058 – 3066.
[588] D. D. Hu, M. R. Eftink, *Biophys. Chem.* **1994**, *49*, 233 – 239.
[589] B. W. Sigurskjold, C. R. Berland, B. Svensson, *Biochemistry* **1994**, *33*, 10191 – 10199.
[590] K. P. Murphy, D. Xie, K. C. Garcia, L. M. Amzel, E. Freire, *Proteins* **1993**, *15*, 113 – 120.
[591] J. Gomez, E. Freire, *J. Mol. Biol.* **1995**, *252*, 337 – 350.
[592] G. C. Kresheck, L. B. Vitello, J. E. Erman, *Biochemistry* **1995**, *34*, 8398 – 8405.
[593] J. J. Christensen, L. D. Hansen, R. M. Izatt, *Handbook of Proton Ionization Heats and Related Thermodynamic Quantities*, Wiley, New York, **1976**.
[594] I. Jelesarov, H. R. Bosshard, *J. Mol. Recognit.* **1999**, *12*, 3 – 18.
[595] K. P. Murphy, E. Freire, *Adv. Protein Chem.* **1992**, *43*, 313 – 361.
[596] R. S. Spolar, J. R. Livingstone, M. T. Record, Jr., *Biochemistry* **1992**, *31*, 3947 – 3955.
[597] P. R. Connelly, J. A. Thomson, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 4781 – 4785.
[598] G. I. Makhatadze, P. L. Privalov, *Adv. Protein Chem.* **1995**, *47*, 307 – 425.
[599] B. M. Baker, K. P. Murphy, *J. Mol. Biol.* **1997**, *268*, 557 – 569.
[600] C. McNemar, M. E. Snow, W. T. Windsor, A. Prongay, P. Mui, R. Zhang, J. Durkin, H. V. Le, P. C. Weber, *Biochemistry* **1997**, *36*, 10006 – 10014.
[601] N. J. Faergeman, B. W. Sigurskjold, B. B. Kragelund, K. V. Andersen, J. Knudsen, *Biochemistry* **1996**, *35*, 14118 – 14126.
[602] G. A. Holdgate, A. Tunnicliffe, W. H. Ward, S. A. Weston, G. Rosenbrock, P. T. Barth, I. W. Taylor, R. A. Pauptit, D. Timms, *Biochemistry* **1997**, *36*, 9663 – 9673.
[603] R. Perozzo, I. Jelesarov, H. R. Bosshard, G. Folkers, L. Scapozza, *J. Biol. Chem.* **2000**, *275*, 16139 – 16145.
[604] M. V. Botuyan, D. A. Keire, C. Kroen, D. G. Gorenstein, *Biochemistry* **1993**, *32*, 6863 – 6874.

[605] J. E. Ladbury, J. G. Wright, J. M. Sturtevant, P. B. Sigler, *J. Mol. Biol.* **1994**, *238*, 669 – 681.
[606] H. Berglund, M. Wolf-Watz, T. Lundback, S. van den Berg, T. Hard, *Biochemistry* **1997**, *36*, 11 188 – 11 197.
[607] F. Dullweber, M. T. Stubbs, J. Stürzebecher, D. Musil, G. Klebe, *J. Mol. Biol.* **2001**, *313*, 593 – 614.
[608] J. R. Liggins, P. L. Privalov, *Proteins* **2000**, *Suppl.*, 50 – 62.
[609] P. L. Privalov, S. A. Potekhin, *Methods Enzymol.* **1986**, *131*, 4 – 51.
[610] K. Takahashi, H. Fukada, *Biochemistry* **1985**, *24*, 297 – 300.
[611] A. Shrake, P. D. Ross, *J. Biol. Chem.* **1990**, *265*, 5055 – 5059.
[612] L. N. Lin, A. B. Mason, R. C. Woodworth, J. F. Brandts, *Biochemistry* **1994**, *33*, 1881 – 1888.
[613] F. Conejero-Lara, P. L. Mateo, *Biochemistry* **1996**, *35*, 3477 – 3486.
[614] K. G. Davis, S. E. Plyte, S. R. Robertson, A. Cooper, G. G. Kneale, *Biochemistry* **1995**, *34*, 148 – 154.
[615] S. V. Litvinovich, K. C. Ingham, *J. Mol. Biol.* **1995**, *248*, 611 – 626.
[616] V. V. Filimonov, V. V. Rogov, *J. Mol. Biol.* **1996**, *255*, 767 – 777.
[617] S. Makino, I. D. Kuntz, *J. Comput. Chem.* **1997**, *18*, 1812 – 1825.
[618] S. Grüneberg, B. Wendt, G. Klebe, *Angew. Chem.* **2001**, *113*, 404 – 408; *Angew. Chem. Int. Ed.* **2001**, *40*, 389 – 393.
[619] H. J. Böhm, *J. Comput. Aided Mol. Des.* **1992**, *6*, 593 – 606.
[620] H. J. Böhm, S. Brode, U. Hesse, G. Klebe, *Chem. Eur. J.* **1996**, *2*, 1509 – 1513.